Adult language learning after minimal exposure to an unknown natural language

Gullberg, Marianne; Roberts, Leah; Dimroth, Christine; Veroude, Kim; Indefrey, Peter

Published in:
Language Learning

DOI:
10.1111/j.1467-9922.2010.00598.x

2010

Adult Language Learning After Minimal Exposure To An Unknown Natural Language

Marianne Gullberg[1,2], Leah Roberts[1], Christine Dimroth[1], Kim Veroude[3], Peter Indefrey[3,4]

[1] Max Planck Institute for Psycholinguistics

[2] Centre for Languages and Literature, Lund University

[3] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour

[4] Institut für Sprache und Information, Heinrich Heine University Düsseldorf

Corresponding author:
Marianne Gullberg
Centre for Languages and Literature
Lund University
PO Box 201
S.221 00 Lund
Sweden
Email: marianne.gullberg@ling.lu.se

Abstract

Despite the literatures on the role of input in adult second language (L2) acquisition, and on artificial and statistical language learning, surprisingly little is known about how adults break into a new language in the wild. This paper reports on a series of behavioral and neuroimaging studies that examine what linguistic information adults can extract from naturalistic but controlled audio-visual input in an unknown and typologically distant L2 after minimal exposure (7 - 14 minutes) without instruction or training. We tested the step-wise development of segmental, phonotactic and lexical knowledge in Dutch adults after minimal exposure to Mandarin Chinese, and the role of item frequency, speech-associated gestures, and word length at the earliest stages of learning. In an exploratory neural connectivity study we further examined the neural correlates of word recognition in a new language, identifying brain regions whose connectivity was related to performance both before and after learning. While emphasizing the complexity of the learning task, the results suggest that the adult learning mechanism is more powerful than is normally assumed when faced with small amounts of complex, continuous audio-visual language input.

Adult Language Learning After Minimal Exposure To An Unknown Natural Language.

"Far too little empirical attention has been paid to the very beginnings of the acquisition process." (Perdue, 1996: 138)

It is a challenging task to learn a new, second language (L2) as an adult outside a classroom and without any help or instruction. Under such circumstances learners must rely entirely on their own capacities for dealing with the language input they are exposed to. The success of such learning depends on the kinds of prior linguistic and nonlinguistic knowledge learners bring to the task, and on the structure of the input. The process of breaking into a new language, what Klein (1986: 59) calls the learner's 'problem of analysis', consists of at least three elements: the segmentation of the continuous speech stream to identify relevant strings such as words; the identification of relevant meaning in the environment that can be mapped onto the sound strings identified, and finally, the generalization beyond exemplars in the input to novel items and the formation of linguistic categories and regularities ('rules'). Despite the considerable literature on adult L2 acquisition, surprisingly little is known about how adult untutored L2 learners go about this complex task 'in the wild'.

The L2 literature has long debated the role of input for learning and what adult learners are able to do with the information available in the linguistic input (see Carroll, 1999, 2001; 2004 for discussions of this notion). The debate is partly prompted by the observation that adult learners do not replicate the information in the input very well, and therefore do not seem to use the available information efficiently. Consequently, a wealth of research has examined possible differences between input and intake (e.g., Schmidt, 1990), the role of attention and noticing (e.g., Gass, Svetics, & Lemelin, 2003;

Izumi, 2002; Robinson, 2003; Schmidt, 2001; Wong, 2001), differences between explicit vs. implicit learning, and declarative vs. procedural knowledge (e.g., Anderson, 1976; DeKeyser, 2003; Ellis, 1994; Hulstijn, 2003; Paradis, 2009; Ullman, 2001), etc.

Interestingly, the theorizing about what the adult L2 learner is or is not capable of both in terms of developing representations and processing, typically draws on data from stages of learning where a considerable amount of knowledge has already been acquired (this is true even for studies explicitly concerned with 'the initial state' of L2 acquisition, see, e.g., papers in Schwartz & Eubank, 1996). That is, theories generally consider learners who already have vocabularies and grammatical systems available to bootstrap the learning of more material in the L2. With a few notable exceptions (e.g., Klein & Dimroth, 2009), most theories of the capacity of the adult learning mechanism are based on intermediate learning stages where knowledge has been acquired which can boost further input processing. This is all the more surprising since the conclusions that can be drawn from language acquisition research about native and non-native speech processing and about the nature of learning itself are dependent on how we see the interaction between the language faculty (which kind of knowledge is relevant, etc.) and the structure of the language input it has to deal with. "[T]he topic remains one of the most under-theorized and under-researched areas of our field." [Carroll, 1999 #1674: 338].

To gauge the limitations and capacities of the adult language learning mechanism it is important to capture it at work at first contact and after minimal exposure to the new language, that is, to control the effects of pre-existing (linguistic) knowledge in order to minimize the effects of previously learned languages (cf., research on cross-linguistic influence in L2; e.g., Jarvis & Pavlenko, 2008). Moreover, to understand which auditory and/or visual features in the input are noticed, attended to and taken as evidence of linguistic distinctions by learners, it is important to control the incoming string and its properties. Studies of artificial language learning (for examples, see this volume) do precisely that but often from a different perspective than L2 studies.

The aim of the studies summarized here has been to achieve these goals while simulating the complexity of the learning task in the wild by exposing adult learners to natural language in an audio-visual (but non-interactive) setting, controlling pre-existing linguistic knowledge and controlling the properties of the input language. Inspired by an unpublished pilot project conducted at the MPI for Psycholinguistics (Zwitserlood et al., 1994), we have developed and successfully used a test paradigm allowing us to examine the earliest perception and processing of input in an unknown natural L2, and the step-wise development of segmental, lexical, and phonotactic knowledge of a new language. We have probed the development both at the behavioral and at the neurological level to examine the capacity to isolate strings or identify forms in the input; to identify relevant meaning from the context and map it onto new sound strings or word forms; and to generalize beyond the input and extract more abstract information such as phonotactic knowledge.

This series of studies differs from related research in its focus on natural, complex and continuous audiovisual input rather than artificial auditory or written input; and in its in the emphasis on implicit learning outside a classroom context; and in the absence of training. We also examine the effects of minutes rather than hours or weeks of exposure. Moreover, in contrast to many studies examining the role of frequency in processing and language learning (for an overview of L2 effects, see Ellis, 2002), we investigate the effect of item frequency at very low limits of frequency. Finally, we target broader contextual cues like speech-associated gestures accompanying speech.

The Input: A Weather Report in Mandarin Chinese

The aim was to test adult learners' capacity to extract linguistic knowledge from textually coherent linguistic input in a natural but unknown language without help while 1) controlling for and minimizing the influence of pre-existing linguistic knowledge, keeping knowledge of known languages constant and knowledge of the target language to zero; and 2) controlling for frequency of a set of target words, and for 'highlighting' of target words in the form of accompanying gestures.

To that end, we constructed seven minutes of controlled but naturalistic input in the form of a weather report in Mandarin Chinese, a language typologically and genetically unrelated to the participants' native language (Dutch) or any language known to them. The weather report consisted of 120 clauses of coherent text containing one out of a set of target word per clause (for details, see Gullberg, Roberts, & Dimroth, accepted). The target words were distributed across the clauses to appear in clause-initial, medial or final position to avoid position effects. The target words were either frequent (8 occurrences) or infrequent (2 occurrences). Half of the target words were also highlighted with gestures forming a deictic link to the referential content, i.e. the <u>icons</u> on the accompanying weather charts (six in total). All other words in the weather report ('padding' words) were also controlled for frequency such that no padding word was more frequent than the frequent target words. Overall, the text consisted of 292 word types (mean number of syllables/clause 7.85).

The highlighting gestures were scripted and to ensure a reliable deictic link between gesture, speech and weather charts, the temporal and spatial accuracy of the gestures was controlled by frame-by-frame analysis of digital video allowing for 40 ms accuracy.

The weather report was presented audio-visually by a female native speaker of Mandarin Chinese who read the text in Chinese characters from a tele-prompter while gesturing as directed. In all the experiments, we presented the weather report to participants prior to the experimental session proper. To ensure that participants did not watch it strategically for a particular task, they were all given the general instruction to 'watch the film'. Throughout, we tested native speakers of Dutch with no knowledge of Mandarin Chinese, as ascertained through an extensive language background questionnaire (Gullberg & Indefrey, 2003).

The First 7-14 Minutes of Contact with an Unknown Language

Word Recognition after 7 and 14 Minutes

Segmenting input in a new language is a crucial first step in language acquisition. This is a challenging task, especially in situations of untutored language learning. Words do not come with predefined breaks between them but learners must rely on cues in the input to detect word forms. Furthermore, for adult L2 learners, it is likely that cues acquired in the first language influence analysis of the incoming string in the second language (transfer or cross-linguistic influence). The segmentation problem can obviously be facilitated in tutored situations where teachers or native speakers may adjust their articulation to emphasize word boundaries through so-called teacher or foreigner talk (e.g., Ferguson, 1975). Mostly, however, learners must identify word strings on their own. The difficulties this causes are manifest in learner-typical behavior such as the production of chunks or formulae (for an overview, see Wray, 2009) where word boundaries are not necessarily observed at early stages (e.g., words in early French learner language such as [levolur] 'steal/thief', [lepeje] 'money/pay' (Perdue, 2006: 860) with unanalyzed 'prefixes' reminiscent of articles). However, previous studies have shown that adult L2 learners draw on fine-grained acoustic cues to detect word boundaries and segment auditory L2 input using both language-specific phonetic and phonotactic information from the L1 (Broersma, 2005; Cutler, 2001; Cutler, Mehler, Norris, & Segui, 1986; Cutler & Otake, 1994; Flege & Wang, 1990; Weber & Cutler, 2006) and more general acoustic information such as aspiration (Altenberg, 2005; Barcroft & Sommers, 2005) to do so. Studies in the field of artificial language learning also suggest that L2 learners are able to use statistical transition probabilities between syllables to determine likely word boundaries (e.g., Saffran, Newport, & Aslin, 1996; Weiss, Gerfen, & Mitchel, 2009, and papers in this volume). Studies of the types cited provide important evidence on learners' strategies and use of cues. However, they leave open the question of what learners are capable of at the very outset of learning, especially at first contact with a new natural language presented as continuous speech.

We probed native Dutch speakers on word recognition in Mandarin Chinese after 7 and 14 minutes of contact (Gullberg et al., accepted). The learners, who had no knowledge of Mandarin Chinese, watched the Chinese weather report and in a subsequent surprise word recognition task had to decide whether sounds played had been heard before. In contrast to studies targeting fine-grained acoustic cues, we examined the effect of gestural highlighting (presences vs. absence) on participants' performance on the assumption that the presence of a gesture might increase the saliency of a string contributing to improved segmentation. We also investigated the effect of word frequency (8 vs. 2 occurrences), and word length defined as number of syllables (one vs. two). Number of syllables was chosen because it has previously been shown that monosyllabic items may cause difficulty of perception in certain contexts (e.g., Dommergues & Segui, 1989).[1] Finally, we examined the effect of amount of exposure, comparing performance after 7 vs. 14 minutes of input.

The results showed that after 7 minutes of Chinese input participants more accurately recognized frequent than infrequent words, and were also more accurate on disyllabic than on monosyllabic words. Participants in fact showed a no-bias towards monosyllabic words, consistently rejecting them. Moreover, word-internal backwards transition probabilities (TPs; the probability of x | y = freq xy / freq y) affected accuracy but only for frequent disyllabic words occurring eight times. That is, the higher the word-internal TP and the more frequent the word in the weather report, the better participants recognized it. In contrast, word-external TPs had no effect on the accuracy scores, providing a possible explanation for why monosyllabic words were poorly recognized and rejected. That is, because monosyllabic words by definition only have word-external TPs, they are harder to recognize. Interestingly, there was no difference between participants who had had 7 vs. 14 minutes of exposure. Finally, the presence of gestural highlighting had no effect on word recognition at these levels of exposure.

These findings overall suggest that adult native Dutch speakers with no prior exposure to Mandarin Chinese can segment the Mandarin sound stream leading them to correctly recognize a disyllabic word that has occurred as little as 8 times in the continuous auditory speech input when it is subsequently presented in isolation. The frequency effect is interesting in that it stresses the difference between very small increments. Four occurrences in sustained speech were not enough for word recognition to take place (low frequent in the double-exposure group) but 8 occurrences were (frequent in the single-exposure group); there was a numerical improvement between 6 and 16 occurrences, although it was not statistically robust. A study testing word recognition in native Dutch speakers and English speakers with no knowledge of Dutch similarly found that both native and non-native listeners recognized words repeated 10 times in isolation but non-native listeners did not recognize words repeated four times in continuous speech in sentence context (Snijders, Kooijman, Cutler, & Hagoort, 2007). Additions in small increments can thus make a big difference to non-native segmentation skills.

Sound to Meaning Mapping after 7 and 14 Minutes

The mapping problem, that is, the linking of meaning to identified word forms, is a vital part of acquisition and one which has received comparatively more attention than the segmentation problem. In the child language literature, word learning, lexical and vocabulary acquisition is often discussed in terms of 'fast mapping' (Carey & Bartlett, 1978) whereby children supposedly 'learn' a new word after a single or very few encounters in the input, typically during the so called vocabulary explosion around 18-24 months (e.g., Clark, 2003; for a discussion of what 'learning' might mean, see Bloom, 2004). In L2 studies, it is generally recognized that adults are competent vocabulary learners, but they are rarely granted the capacity for fast mapping.

A large body of literature investigates adult L2 learners' acquisition of productive and receptive vocabulary, the role of consciousness and attention for lexical acquisition, and implicit and incidental word learning (e.g., Bogaards & Laufer, 2004; DeKeyser, 2003 for overviews; Ellis, 1994; Hulstijn, 2001; 2003). Many studies focus on word learning through reading, showing that adults can learn new words without instruction while reading for comprehension after two to three encounters (e.g., Horst, Cobb, & Meara, 1998; Hulstijn, Hollander, & Greidanus, 1996; Rott, 1999). There are frequency effects such that the number of repetitions often improve learning (Kirsner, 1994), but so do cognate status and the size of the pre-existing vocabulary (see Hulstijn, 2003 for an overview). As in the case of the segmentation problem, very little is known about adults' capacity for vocabulary acquisition at the outset when no existing L2 vocabulary constrains learning of novel items, and when words are presented in continuous speech without any didactic intent. Whether or not adults can fast map under such circumstances is not known (see Rohde & Tiefenthal, 2000 for a study on fast mapping with training).

In another set of experiments (Gullberg et al., accepted), we examined whether adult native speakers of Dutch can extract meaning from Mandarin Chinese input and map it onto sound strings from the input, what is, whether they can map a sound string to a target icon from the weather report. A new set of native speakers of Dutch with no knowledge of Mandarin Chinese watched the weather report and then participated in a surprise auditory picture-word matching task, where they had to decide whether sounds matched a weather icon shown on screen. The experiment focused on the target nouns from the weather report (e.g., sun, cloud). Again, we examined the effect of word frequency, gestural highlighting, word length, TPs and amount of exposure on participants' performance. The results indicated that participants were significantly more accurate on disyllabic items that had been frequent and gesturally highlighted in the weather report. As was the case for word recognition, word-internal backwards TPs had an effect on accuracy such that the higher the word-internal TP and thus the more frequent the word, the more accurately the word was paired with the appropriate weather icon.

Again, the findings suggest that adult native Dutch speakers with no prior exposure to Mandarin Chinese can map meaning to disyllabic strings in the Mandarin sound stream that have occurred as little as 8 times in the input in combination with a gesture to form a link to the referential content on the weather charts. The effect of gestural highlighting is not surprising and tallies well with studies showing that viewers integrate the information conveyed by gestures to improve comprehension (e.g., Beattie & Shovelton, 1999; Butterworth & Itakura, 2000; Langton, O'Malley, & Bruce, 1996). Studies in other domains also suggest that contextual, visual cues such as the speaker's face and mouth improve lexical learning (Davis & Kim, 2001; Reisberg, McLean, & Goldfield, 1987). More newsworthy is the observation that the mapping of meaning to word form at these earliest stages seems to require accumulative cues to take full effect such that gestures and frequency must work together.

Extracting Regularities And Generalizing: Phonotactic Information After 7 To 14 Minutes

In addition to lexical learning, acquisition crucially involves the capacity to generalize away from the input and encountered exemplars to the formation of categories and the establishment of regularities or 'rules'. Traditional L2 studies have given a lot of attention to the (longitudinal) development of regularities in production, especially in the domain of morphosyntax, arguing for developmental sequences and stages typical of all interlanguage and at least partly independent of source and target language structures (e.g., Klein & Perdue, 1997; Meisel, Clahsen, & Pienemann, 1981 inter multa alia). In the domain of comprehension, studies of artificial language learning in turn typically show that adults are capable of using statistical and prosodic information such as pauses for morphosyntactic rule formation even of non-adjacent and nested types after exposure to constrained artificial languages (e.g., de Diego Balaguer & Barroso; Folia et al., this volume). Again, despite all these efforts, it remains unclear whether adults can detect abstract regularities in the input and generalize such regularities to novel items after minimal exposure to natural language (for morphological learning after longer exposure to natural language, see Davidson, this volume).

We were interested in whether adults can extract phonotactic information, that is, highly abstract information about the sound structures, from continuous natural language input in a new language. It has been suggested that phonotactic acquisition shows frequency and statistical effects such that novel words with low transition probabilities between segments are judged as non-words compared to words with relatively higher segmental TPs which are instead judged as pseudo words, that is, as possible words in the language (Frisch, Large, Zawaydeh, & Pisoni, 2001).

We tested whether Dutch adults could detect syllable structure violations in Mandarin Chinese after minimal exposure, and whether they could apply phonotactic knowledge derived from the input to new items of the language (Roberts, Dimroth, & Gullberg, in prep.). As before, participants watched the weather report and then completed a surprise lexical decision task where they listened to sounds and had to determine whether they were ‘real Chinese’ or not. We tested two groups with 7 vs. 14 minutes of exposure to the weather report. We also tested a third control group with no input at all in order to control for the fact that even Dutch speakers with no formal knowledge of Mandarin Chinese have a preconceived idea about what Mandarin sounds like. The experimental materials consisted of real monosyllabic words, half of which had appeared in the weather report (e.g., *yun2* ‘cloud’) and half of which were new, and monosyllabic words containing phonotactic violations. A set of filler items had violations consisting of three- and two-consonant clusters word-finally (e.g., *alst, ans*) or word-initially (e.g., *spra, sna*). These Germanic-sounding cluster violations were assumed to be easy to reject for Dutch speakers as not being Mandarin. The experimental test items consisted of CVC syllables ending in an illegal word-final consonant in Mandarin (e.g., *gam*).

The results showed that all three groups correctly rejected all the filler three-consonant cluster and word-final two-consonant cluster syllables as not being Mandarin Chinese even without exposure to the weather report. For word-initial 2-consonant clusters, everybody was at chance. Most strikingly, judgments about the experimental CVC violation syllables were at chance for the control group, but with increasing exposure participants more accurately rejected the illegal CVC syllables. In other words, even without exposure Dutch speakers know something about both Mandarin Chinese and Dutch phonotactic structure, but crucially, they are also able to draw on minimal and complex natural language input to extract information about the sound structure of the new language. Importantly, the ability to identify illegal CVC syllables in Chinese must stem from an analysis of the new language input and cannot be based on transfer of L1 distinctions, since CVC syllables of this type (e.g., *gam*) are acceptable in the source language Dutch. These results provide evidence not only for an ability to roughly segment and recognize items previously encountered, but for an ability to generalize phonotactic knowledge to new items after as little as 7 minutes of contact.

Neural Correlates of Word Recognition

To complement the behavioral studies, we have also examined the neural correlates of the first minutes of learning of a new language. To date, most neuroimaging studies investigating the neural correlates of learning new words have presented participants with isolated novel word forms (Breitenstein et al., 2005; Cornelissen et al., 2004; Davis, Di Betta, Macdonald, & Gaskell, 2009; Grönholm, Rinne, Vorobyev, & Laine, 2005; Mei et al., 2008; Raboyeau et al., 2004; Wong, Perrachione, & Parrish, 2007), often paired with pictures of novel or familiar objects providing the meaning of the novel words. Although there is only partial overlap between the brain regions reported to be activated in these studies, some regions, such as the left inferior frontal cortex, the premotor cortex, the bilateral inferior parietal cortex, in particular the supramarginal gyri, the insula, the left posterior temporal cortex, and the hippocampus have been consistently found to be involved in word learning.

Naturalistic L2 exposure shares with these studies the aspect of the perception and storage of novel word forms (if successfully segmented) but differs with respect to the additional demands due to the segmentation of possible words out of the continuous speech. We were therefore interested to see which of the previously observed regions would also be activated during the processing of naturalistic continuous speech in an unknown language and hence be related to successful word form segmentation. Unlike the sequential presentation of novel words used in previous studies, however, blood flow changes measured during the exposure to weather report video clips are not well suited for standard functional magnetic resonance (fMRI) analyses, because they are based on a statistical model that takes into account the temporal order and duration of conditions. In a recent study (Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010) we therefore used a model-free approach, assessing the so-called functional connectivity, that is, correlations between the fluctuations of hemodynamic activation of spatially distinct areas (Friston, Frith, Liddle, & Frackowiak, 1993). During rest, functionally related brain regions display correlations in the fMRI time courses (Biswal, Yetkin, Haughton, & Hyde, 1995) and temporary changes in correlations between brain regions can be the result of task performance. Waites, Stanislavsky, Abbott, and Jackson (2005), for example, found an increase in the correlations between regions involved in a language task in resting state after performance of the task.

In our study we assessed the functional connectivity of the regions previously reported to be involved in word learning during resting periods of five minutes before, between, and after two presentations of the Mandarin Chinese weather report. In addition to motor-related regions (supplementary motor area, SMA and insula) which typically show strong interhemispheric connectivity, we found the left and right supramarginal gyri to show increasing functional connectivity over time reaching highest connectivity during the second run of the weather report movie and the last resting state period. Furthermore, during the last resting state period the connectivity between the supramarginal gyri was stronger for a sub-group of participants who showed some ability to recognize Chinese words ('learners') compared to a sub-group who performed at chance level on a post-experiment word recognition test ('non-learners').

Taken together, these findings suggest an involvement of the supramarginal gyri in the successful segmentation and storage of phonological representations of Chinese words. This interpretation is supported by an fMRI study by McNealy, Mazziotta, and Dapretto (2006) who used an artificial-language paradigm to study the neural correlates of the ability of listeners to segment 'words' out of streams of CV syllables based on statistical cues, such as different frequencies of syllable co-occurrence. These authors found that hemodynamic activation in bilateral posterior temporal and inferior parietal regions including the supramarginal gyri increased more strongly as a function of exposure duration when the language stream provided cues to word segmentation than when it did not.

Whereas the supramarginal gyri only showed connectivity differences between learners and non-learners after exposure to the weather report, we found two other region pairs (left insula and Rolandic operculum as well as left SMA and precentral gyrus) to show stronger connectivity for learners only before exposure to the weather report. These regions are known to be involved in articulation and phonological rehearsal. One somewhat speculative interpretation might thus be that learners and non-learners differed with respect to a predisposition to involve the speech motor system during perception.

Discussion and Conclusions

The findings from the experiments reviewed here suggest that adult learners are able to deal very efficiently and quickly with very complex input even in the absence of instructions. They are able to extract segmental, word-form related information, lexical meaning from the context and map it onto word forms identified, and finally, to extract abstract, phonotactic information and generalize it to novel items not encountered in the input after as little as seven minutes of contact with an unknown language. This is a remarkable feat. The results complement findings from statistical language learning studies indicating that transition probabilities (TPs) between syllables help learners identify words. However, given the statistical properties of the naturalistic input in this study, with much higher word type counts and lower TPs than is typical in artificial language learning studies, it is perhaps not surprising to find that TPs within words are the only relevant ones, and moreover, that learners draw on both TPs and item frequencies at the levels of exposure we are investigating here. That is to say, there is a powerful statistical reckoning mechanism at work that takes into account both the micro regularities provided by TPs and the coarser statistics provided by whole word forms. This seems to be an efficient solution for dealing with the messier input that is typical of natural languages.

Further, to solve the task of identifying contextual meaning and mapping it onto word strings, learners combine gestural deictic links between icons and word forms with TPs and item frequency to home in on relevant form-meaning pairs. Again, learners exploit all available information in remarkably efficient fashion but it is noteworthy that cues need to accumulate to take effect at these levels of exposure. That is to say, gestures on their own are not sufficient for meaning mapping, as is often assumed in the literature considering the attention directing properties of gesture for first language acquisition (e.g., O'Neill, Bard, Linnell, & Fluck, 2005).

Importantly, the findings from this study complement existing studies of implicit learning, showing that adults can extract both form- and meaning-related information from sustained speech even in the absence of conscious learning efforts. It is important to be able to show that learners can do this at the outset of learning even when they do not have vocabularies and grammar to guide further learning. It is also important to show that the same mechanisms posited in artificial language learning studies work when the input consists of complex natural language. It has been suggested in the traditional acquisition literature that the quest for meaning drives acquisition (e.g., VanPatten, 2002), an assumption supported by the robust finding that content words are acquired before function words by both child and adult learners, by tutored and untutored learners alike (e.g., Clark, 2003; Klein, 1986; Kotsinas, 1983). However, while the quest for content may be a primary *conscious* driving force, segmentation must precede it (see Carroll, 2001 for the same argument) and that must happen in the absence of identified meaning. The word recognition data from this study suggest that such learning is possible.

The neurocognitive study, drawing on model-free fMRI analysis, which allows for the study of neural structures involved in the processing of naturalistic audiovisual L2 input, provides important evidence that there are both pre-existing and learning-induced neural differences in the supramarginal gyri between learners who are more and less successful at word recognition observable after no more than 14 minutes of exposure. An important implication for neurocognitive studies of language learning is that the supramarginal gyri thus appear to be implicated in creating phonological representations of possible L2 words after short exposure to a new language. The findings also have two theoretical implications for L2 studies. First, the L2 literature on age effects or maturational constraints on L2 acquisition has often offered neurological arguments to account for adults' less successful L2 acquisition, mainly in production, appealing to loss of plasticity in the adult brain (for overviews of such arguments, see e.g., Birdsong, 2006; DeKeyser & Larson-Hall, 2005; Singleton & Ryan, 2004). This view of the adult learner as neurologically inflexible clearly has to be modified in view of findings from comprehension studies indicating rapid neurological adjustment to new input after as little as 14 minutes (see also other papers in this volume).

Second, the observation that there are pre-existing neural differences between more or less successful learners is relevant to the literature on individual differences and the thorny issue of language learning aptitude (e.g., Dörnyei, 2005). More specifically, the findings suggest that one cause of individual differences in the word segmentation ability might be related to the recruitment of the speech motor system during perception.

The experiments presented here are the starting point for a range of further investigations. The question of possible maturational constraints on L2 acquisition – both in production and in comprehension – remains a key issue for language acquisition studies. While the literature has often focused on end states, recent studies comparing child L2 learners of different ages suggest that younger children progress via a different route towards the target language than older children (e.g., Dimroth, 2008). Furthermore, age interacts with the linguistic domain under study. For instance, adults are faster or at least not slower than children in initial stages of acquisition for morphosyntax (e.g., Slavoff & Johnson, 1995) or phonology (e.g., Loewenthal & Bull, 1984). In the lexical domain, in contrast, children are sometimes thought to be both faster and better than adults (Carey & Bartlett, 1978). However, Markson & Bloom (1997) found that adults were significantly better than three- and four-year-olds at single word retention when tested immediately after training but that their advantage disappeared when tested a week after exposure. This suggests that rate and degree of retention need to be treated separately. And again, most studies of this type have not examined the very initial stages. We have begun to probe these issues by comparing word recognition and phonotactic learning in six-year-olds and adults (Roberts et al., in prep.).

A further issue to examine concerns how adults' initially efficient processing of new language input tallies with their observed slower progress once the system becomes more complex. To study changes in the rate and degree of retention requires carefully crafted longitudinal studies which remains an important challenge for the future.

A related issue is to study the relationship between the earliest skills in reception and production in order to elucidate how the two modes of language use may interact in acquisition.

In conclusion, the studies presented here suggest that at the earliest stages of L2 acquisition and in the absence of pre-existing knowledge to bootstrap and boost learning, the adult learning mechanism can deal efficiently with very little, and very complex input. The adult learning mechanism appears to be considerably more powerful than typically assumed in the L2 acquisition literature. The combination of statistical properties, frequency, contextual cues and a limited search domain such as the weather report all provide a powerful scaffolding system allowing the adult learning mechanism to extract considerable amounts of linguistic information implicitly. Although above-chance performance on experimental tasks is a far cry from successful second language acquisition, we believe we have made some progress towards answering Clive Perdue's call for more empirical research on the earliest stages of acquisition which is crucial to our understanding of L2 acquisition.

Author Note

We gratefully acknowledge support from the Max Planck Institute for Psycholinguistics, and funding from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO, MPI 56-384, The Dynamics of Multilingual Processing, awarded to M. Gullberg, P. Indefrey, and W. Klein. We also thank Kim Veroude, Aoju Chen, Jidong Chen, Rosa Zhang, Arna van Doorn, Nanna Haug Hilton, Wilma Jongejan, Gerrit-Jan Kootstra, and Linda van Meel. Correspondence concerning this article should be addressed to Marianne Gullberg, Centre for Languages and Literature, Lund University, PO Box 201, S.221 00 Lund, Sweden. Internet: marianne.gullberg@ling.lu.se.

Notes

[1] Artificial language learning studies suggest that transition probabilities (TPs) between syllables are an important factor in learning (e.g., Aslin, Saffran, & Newport, 1998; Perruchet & Desaulty, 2008). However, in natural language material like the weather report, which displays a high number of word types, TPs are quite different from the minimal systems used in artificial language studies. To be able to compare possible effects of item frequency and TPs, we computed TPs for the weather report and as expected, found that word-internal TPs were higher than word-external TPs meaning that words of two syllables may be better recognized than monosyllabic items.

References

Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research*, *21*, 325-358.

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Erlbaum.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 27*, 321-324.

Barcroft, J., & Sommers, M. S. (2005). Effects of acoustic variability on second language vocabulary learning. *Studies in Second Language Acquisition, 27*, 387-414.

Beattie, G., & Shovelton, H. (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? *Semiotica, 123*, 1-30.

Birdsong, D. (2006). Age and second language acquisition: A selective overview. *Language Learning, 56*, 9-49.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*, 537-541.

Bloom, P. (2004). Myths of word learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 205-224). Cambridge, MA: MIT Press.

Bogaards, P., & Laufer, B. (Eds.). (2004). *Vocabulary in a second language: Selection, acquisition and testing.* Amsterdam: Benjamins.

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A.-F., Sommer, J., Wolbers, T., et al. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage, 25*, 958-968.

Broersma, M. (2005). *Phonetic and lexical processing in a second language.* Unpublished Ph.D., Radboud University, Nijmegen.

Butterworth, G., & Itakura, S. (2000). How the eyes, head and hand serve definite reference. *British Journal of Developmental Psychology, 18*, 25-50.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17-29.

Carroll, S. E. (1999). Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning, 49*, 37-92.

Carroll, S. E. (2001). *Input and evidence. The raw material of second language acquisition.* Amsterdam: Benjamins.

Carroll, S. E. (2004). Segmentation: Learning how to 'hear' words in the L2 speech stream. *Transactions of the Philological Society, 102*, 227-254.

Clark, E. V. (2003). *First language acquisition.* Cambridge: Cambridge University Press.

Cornelissen, K., Laine, M., Renvall, K., Saarinen, T., Martin, N., & Salmelin, R. (2004). Learning new names for new objects: Cortical effects as measured by magnetoencephalography. *Brain and Language, 89*, 617-622.

Cutler, A. (2001). Listening to a second language through the ears of a first. *Interpreting, 5,* 1-18.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25*, 385-400.

Cutler, A., & Otake, T. (1994). Mora or phoneme–further evidence for language-specific listening. *Journal of Memory and Language, 33*, 824-844.

Davis, C., & Kim, J. (2001). Repeating and remembering foreign language words: Implications for language teaching systems. *Artificial Intelligence Review, 16*, 37-47.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience, 21*, 803-820.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford: Blackwells.

DeKeyser, R. M., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. De Groot (Eds.), *Handbook of bilingualism. Psycholinguistic approaches* (pp. 88-108). Oxford: Oxford University Press.

Dimroth, C. (2008). Age effects on the process of L2 acquisition? Evidence from the acquisition of negation and finiteness in L2 German. *Language Learning, 58*, 117-150.

Dommergues, J.-Y., & Segui, J. (1989). List structure, monotony, and levels of processing. *Journal of Psycholinguistic Research, 18*, 245-253.

Dörnyei, Z. (2005). The psychology of the language learner: Individual differences in second language acquisition. Mahwah, NJ: Erlbaum.

Ellis, N. C. (1994). Implicit and explicit language learning - An overview. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-31). London: Academic Press.

Ellis, N. C. (2002). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition, 24*, 297-339.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics, 17*, 1-14.

Flege, J., & Wang, C. (1990). Native-language phonotactic constraints affect how well Chinese subjects perceive the word-final English /t/-/d/ contrast. *Journal of Phonetics, 17*, 299-315.

Frisch, S. A., Large, N. R., Zawaydeh, B., & Pisoni, D. B. (2001). Emergent phonotactic generalizations in English and Arabic. In J. Bybee & P. Hopper (Eds.), *Proceedings of the symposium on frequency effects and emergent grammar* (pp. 159-180). Amsterdam: Benjamins.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow and Metabolism, 13*, 5-14.

Gass, S. M., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning, 53*, 497-546.

Grönholm, P., Rinne, J. O., Vorobyev, V., & Laine, M. (2005). Naming of newly learned objects: a PET activation study. *Cognitive Brain Research, 25*, 359-371.

Gullberg, M., & Indefrey, P. (2003). *Language background questionnaire.* From the project The Dynamics of Multilingual Processing. Nijmegen: Max Planck Institute for Psycholinguistics. http://www.mpi.nl/research/research-projects/the-dynamics-of-multilingual-processing/tools/Lang-Hist-Quest-Engl.pdf.

Gullberg, M., Roberts, L., & Dimroth, C. (accepted). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics*.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*, 207-223.

Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Oxford: Blackwells.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal, 80*, 327-339.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition, 24*, 541-577.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition.* New York: Routledge.

Kirsner, K. (1994). Implicit processes in second language learning. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 283-312). San Diego, CA: Academic Press.

Klein, W. (1986). *Second language acquisition*. Cambridge: Cambridge University Press.

Klein, W., & Dimroth, C. (2009). Untutored second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 503-522). New York: Academic Press.

Klein, W., & Perdue, C. (1997). The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research, 13*, 301-347.

Kotsinas, U.-B. (1983). On the acquisition of vocabulary in immigrant Swedish. In H. Ringbom (Ed.), *Psycholinguistics and foreign language learning* (pp. 75-100). Åbo: Åbo Akademi University Press.

Langton, S. R. H., O'Malley, C., & Bruce, V. (1996). Actions speak no louder than words: Symmetrical cross-modal interference effects in the processing of verbal and gestural information. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1357-1375.

Loewenthal, K., & Bull, D. (1984). Imitation of foreign sounds: what is the effect of age? *Language and Speech, 27*, 95-97.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature, 385*, 813-815.

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: Neural mechanisms underlying speech parsing. *Journal of Neuroscience, 26*, 7629-7639.

Mei, L., Chen, C., Xue, G., He, Q., Li, T., Xue, F., et al. (2008). Neural predictors of auditory word learning. *Neuroreport, 19*, 215-219.

Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*, 104-135.

O'Neill, M., Bard, K. A., Linnell, M., & Fluck, M. (2005). Maternal gestures with 20-month-old infants in two contexts. *Developmental Science, 8*, 352-359.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: Benjamins.

Perdue, C. (1996). Pre-basic varieties: The first stages of second language acquisition. *Toegepaste taalwetenschap in artikelen* (Special Issue EUROSLA 6. A selection of papers), *2*, 135-149.

Perdue, C. (2006). “Creating language anew”: some remarks on an idea of Bernard Comrie's. *Linguistics, 44*, 853-871.

Perdue, C. (Ed.). (1993). *Adult language acquisition: Cross-linguistic perspectives*. Cambridge: Cambridge University Press.

Perruchet, P., & Desaulty, S. (2008). A role for backward transitional probabilities in word segmentation? *Memory & Cognition, 36*, 1299-1305.

Raboyeau, G., Marie, N., Balduyck, S., Gros, H., Démonet, J.-F., & Cardebat, D. (2004). Lexical learning of the English language: a PET study in healthy French subjects. *NeuroImage, 22*, 1808-1818.

Reisberg, D., McLean, J., & Goldfield, A. (1987). Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli. In R. Campbell & B. Dodd (Eds.), *Hearing by Eye: The psychology of lip-reading* (pp. 97-114). Hillsdale, NJ: Erlbaum.

Roberts, L., Dimroth, C., & Gullberg, M. (in prep.). Investigating what word form knowledge adults and children can acquire after the first few minutes of exposure to a new language.

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). Oxford: Blackwells.

Rohde, A., & Tiefenthal, C. (2000). Fast mapping in early L2 lexical acquisition. *Studia Linguistica, 54*, 167-174.

Rott, S. (1999). The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition, 21*, 589-619.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*, 606–621.

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

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-30). Cambridge: Cambridge University Press.

Schwartz, B. D., & Eubank, L.(Eds.) (1996). What is the 'L2 initial state'? Special Issue *Second Language Research, 12*.

Singleton, D., & Ryan, L. (2004). *Language acquisition: The age factor* (2nd ed.). Clevedon: Multilingual Matters.

Slavoff, G. R., & Johnson, J. S. (1995). The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition, 17*, 1-16.

Snijders, T. M., Kooijman, V., Cutler, A., & Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research, 1178*, 106-113.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105-122.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52*, 755-803.

Veroude, K., Norris, D. G., Shumskaya, E., Gullberg, M., & Indefrey, P. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language, 113*, 21-27.

Waites, A. B., Stanislavsky, A., Abbott, D. F., & Jackson, G. D. (2005). Effect of prior cognitive state on resting state networks measured with functional connectivity. *Human Brain Mapping, 24*, 59-68.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second language listening. *Journal of the Acoustical Society of America, 119*, 597-607.

Weiss, D. J., Gerfen, C., & Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: A challenge for statistical learning? *Language Learning and Development, 5*, 30 - 49.

Wong, P. C., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping, 28*, 995-1006.

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23*, 345-368.

Wray, A. (2009). *Formulaic language: pushing the boundaries*. Oxford: Oxford University Press.