Second Language Research 8, 1 (1992); pp. 39–58

Assignment of gender to French nouns in primary and secondary language: a connectionist model

M.E. Sokolik and **Michael E. Smith** *Texas A&M University*

In French, grammatical gender is often represented phonologically and/or morphologically. Thus, a language learner's competence for gender identification might in part reflect the ability to recognize patterns in noun phonology and morphology. We herein describe a computer-based connectionist-type network model which learned to identify correctly the gender of a set of French nouns. Subsequently, this model was able to generalize from that learning experience and assign gender to previously unstudied nouns with a high degree of reliability. This gender assignment was accomplished by relying solely upon information inherent in the structure of the nouns themselves, and it occurred in the absence of explicit rules for the evaluation of nouns. Instead, the model discovered criterial gender-specific features when shown examples of masculine and feminine nouns during its initial training period. The model's ability to learn these gender-specific features was found to be related both to its initial connectivity state and to a variable learning-rate parameter. These latter results are discussed with respect to their general implications for second language learning.

I Introduction

French nouns, in addition to being marked for number, are marked for gender: either masculine or feminine. Apart from a subset of nouns that reflect semantic gender (e.g., femme, 'woman', or homme, 'man'), and five other limited classes of nouns,[1] it is widely believed that there is no semantic basis for gender assignment in French. The gender of French nouns can be reflected phonologically or morphologically in some cases by fairly regular endings (Stevens, 1984; Tucker *et al.*, 1968), although this is not always the case. Outside of the phonology and morphology of the noun itself, the articles and adjectives used to modify nouns also agree in gender with the gender of the noun, which in the case of irregularly patterned nouns, may help to discriminate the gender of those nouns when other features do not (Surridge, 1986).

[1] The other five classes, according to Valdman (1969: 145) are: 1) most chemical elements and compounds; 2) colour terms in the abstract; 3) names of languages; 4) names of trees; and 5) names of geometric figures.

The acquisition of noun gender association is considered to occur relatively early in native speakers (age 3;0; Clark, 1985), but in the hierarchy of second language acquisition, it is considered a late-learned feature (Andersen, 1984). Given the age of acquisition for first language, it is fairly safe to say that it would be unusual that a native speaker's facility with gender would be attributable to any formal instruction or overt rules for recognizing or using correct gender markings. Similarly, when learning French as a second language, guidelines for assigning gender are only sometimes overtly taught:

> Since no rules for distinguishing gender classes are taught [the English speaker], the whole process seems completely arbitrary. In contrast, French speakers, including very young children, seem to have no difficulty with gender, even with the gender of novel nouns (Tucker, *et al.*, 1968: 312).

Instead, the traditional view is that the second language learner has three possible routes for the acquisition of noun gender association, which are, of course, not mutually exclusive:

1) learn to recognize that certain orthographic or phonetic groupings in nouns are predictive of gender assignment;
2) rely upon contextual information that specifies noun gender, that is, articles and adjective agreement, over the course of vocabulary acquisitions; and
3) rely upon rote memorization of nouns with associated gender-marked articles.

From these three processes, one might suppose that some rules for gender classification are inferred in the course of language acquisition. However, the typical second language learner learning French (as well as many native speakers) may not be able to formulate overtly any such rules in a coherent and consistent fashion (cf. Lewicki *et al.*, 1988; Reber, 1976). An alternative possibility is that much of what is being acquired during the acquisition of gender is an implicit set of pattern detectors rather than an explicit rule system. That is, gender might often be automatically assigned during the perceptual process rather than inferred from available contextual information or from retrieval of associated information in memory.

This possibility raises a few important questions. First, is the information inherent in the structure of individual French nouns sufficient for the reliable assignment of gender by some relatively low-level perceptual process? Secondly, what type of perceptual mechanism could accomplish such a task? Finally, how might such a mechanism evolve out of the language learner's interaction with a given stimulus environment?

In this paper, we assess whether the information inherent within the structure of individual French nouns is sufficient to allow gender to be correctly assigned without reliance upon other types of information. Further, we examine the acquisition of gender-recognition skill in both primary and secondary language.

The formal aspects of this analysis are embodied in a computational 'connectionist' or 'parallel distributed processing' (PDP) model. The architecture of these types of models is inspired in part by brain architecture (Clark, 1989), and they have been demonstrated to exhibit some of the robust pattern recognition capabilities characteristic of human cognitive processes (Rumelhart and McClelland, 1986). Information in connectionist models is codified by a specific pattern of activity distributed over a set of low-level featural nodes or 'units' (in some ways analogous to neurons in the brain). A processing module might be composed of multiple populations or 'layers' of such units. Two given layers can be related by a set of pairwise interconnections (analogous to synapses) of individual units across layers, and activity may be shared between units via these interconnections. The strength or 'weight' of connectivity between specific pairs of units varies, and thus the degree to which activity is shared between pairs of units also varies.

Cognitive transformations in these types of architecture do not rely upon explicit rule systems. Instead, the transformation a given module performs is implemented as a mapping of activity from one layer of units to another layer of units. The form this transformation takes is determined by the pattern of pairwise weights that relate the units at one level to those at the next. For example, the activity of a unit in one layer will have a relatively large positive influence on the activity of a unit in the next layer if those two units are interconnected by a relatively large positive weight, or conversely, a small or negative (inhibitory) influence if the weight is small or negative.

The ability of a given processing module to accomplish anything useful is thus dependent upon the pattern of its interconnection weights. The critical developmental issue is therefore the question of how to determine an appropriate set of weights for a particular desired transformation. One option is to specify the weights a priori, given that an appropriate set of weights to accomplish the transformation is known in advance. However, a more naturalistic approach is to allow the model to evolve a functional set of weights in response to its interaction with a stimulus environment. Thus, in many connectionist models the weights are modified over time in a manner that reflects the processing history of the system. That is, the strengths of connectivity between units in different layers come to reflect learned associations. When equipped with a suitable modification function, such networks have been demonstrated to possess the ability to learn complex mappings between representations. Thus, given only an arbitrary set of pairs of input and output patterns, connectionist networks can often learn subsequently to produce an appropriate output pattern when presented with a particular input pattern. This type of pattern recognition is performed automatically without any 'look-up' procedure or explicit application of 'if-then'-type rules.

The particular type of network analysed in this study is often referred to as a 'pattern associator'. Pattern associators are one of the most basic types of connectionist architectures, where a layer of units representing an input pattern is connected by a single set of modifiable connections to a layer of units representing an output pattern or classification scheme. Connectionist architectures of this kind have been widely studied as learning paradigms for both linguistic and nonlinguistic domains (Anderson, 1983; Kohonen, 1984; Hinton and Anderson, 1981; McClelland and Rumelhart, 1986). Two important properties of pattern association networks are consistent with those minimally required for models of human learning mechanisms: they can learn to abstract features common to a set of inputs; and they can generalize on the basis of those features to permit the categorization of new patterns that are similar to patterns that occurred in a previously studied set.

The computational experiments performed with the pattern associator developed here were designed to accomplish three things. First, at a most basic level, we examined whether a connectionist-style model can successfully evolve a set of connections that enable it to classify French nouns on the basis of their gender association. In subsequent experiments we extended our basic model to conditions more like those presumed to exist for the second language learner (cf. Jacobs, 1988; Sokolik, 1990). That is, we examined the computational effects of pre-existing patterns of connectivity (such as those presumed to be formed during first language acquisition) and decreases in plasticity (such as that presumed to be present in adult language learners) on the ability of the model to acquire gender associations. Finally, we examined whether the network could generalize the knowledge it acquired under these various learning conditions. In other words, could the model reliably classify a set of nouns it had never encountered before, based solely on the learning experience it had with another set of nouns.

II General methods

1 The stimulus set

To create a set of stimulus items representative of those a second-language learner might encounter, all of the vocabulary-item nouns of three to eight letters in length were culled from an introductory French language textbook.[2] This produced a corpus of 406 masculine gender nouns, and 317 feminine gender nouns. Three hundred items were then randomly selected from each of these lists, and the resulting group of 600 (Appendices A and B) words served as the stimulus materials. From this 600-item list, 450 words were randomly selected and intermingled to serve as the set of materials that the model was originally trained on (225 masculine and 225 feminine). The remaining 150 words were used to test the ability of the model to identify the gender of nouns it had not previously encountered. The wordlist contained both 'regular' nouns, like those with common masculine (-eur, -ier, -ment, etc.), and feminine endings (-tion, -euse, -iere, -esse, etc.), as well as many irregular nouns. These nouns could include either an unusual ending, that is, an ending that did not fit into any established pattern (e.g., *faim* 'hunger' (f.)), or a reversal of the expected pattern, such as an -eur ending occurring on a feminine gender noun (e.g., *peur* 'fear' (f.)).

2 Network architecture

In the pattern associator studied here, the input layer of units coded the stimulus items for letter and position (Figure 1). A 28-item alphabet was utilized with items 1–26 coding the letters A–Z, and items 27 and 28 coding è and é, respectively. (Circumflexes were not coded as such, but were coded as the letter *s*, which the circumflex replaces in modern French orthography. For example, the word *fenêtre* 'window' was coded as *fenestre*.) As a result, the input layer (henceforth IL) was a vector of nodes or units of length 224 (8 positions × 28 possible letters for each).

As shown in the figure, the presence of a letter in a position gave the corresponding unit an activation of 1, and its absence an

[2] Pinker and Prince (1988) claim that standard orthography is an unsatisfactory system for input into network architecture. They also present criticism of other systems used (for example, the Wickelphone system). Given the inherent problems with other systems, and the fact that we are not making substantial claims as to comparability with first language acquisition, the argument against using standard orthography becomes weaker. In fact, given that frequently a second language learner *does* rely on the orthographic representations of words, this seems even less a problem. In any case, the major conclusions reached are likely to be independent of the coding scheme utilized.

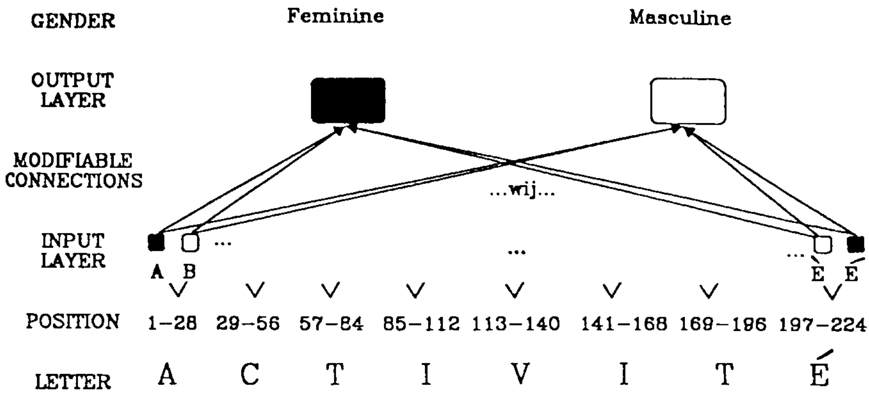

Figure 1 Graphic representation of network architecture. The 224 input nodes code both letter identity and letter position, and the two output nodes code word gender. Each output node is connected to each input node by a variable connection strength that changes as a result of learning.

activation of 0. For example, if a word began with an 'A', the first unit in the vector would have an activation of 1, and units 2 through 28 would have an activation of 0. In the case of words with fewer than eight letters, the word was justified with the rightmost position in the input buffer. For example, for the five-letter word homme, no active unit occurred in the first three sets of input letters. The output layer (henceforth OL) consisted of only two units, one to represent feminine gender, and the other to represent masculine.

The values or activations that these units took on in any given trial was determined by the stimulus pattern and the strength of connectivity between each IL unit and each OL unit. That is, the activation of an OL unit was determined by

$$a_i = \Sigma a_j w_{ij} \tag{1}$$

where a_i represents the activation of an OL unit, a_j represents the activation of an IL unit, and w_{ij} represents the weight or strength of connectivity between them. More informally, the activation of an OL unit is equal to the sum of the activations of the IL units times the strengths of their connections to the OL unit.

3 The network learning procedure

In different learning conditions, the network was either initialized with all of the connections between IL and OL set to zero (that is, it sometimes began as a *tabula rasa*), or instead, its initial connection weights between IL and OL units were randomly prespecified. In either case, the task of the learning procedure was then to determine a set of weights such that, given a feminine input word, a greater activation would be produced on the feminine OL unit than the masculine, and vice versa. One learning procedure that has been successfully employed to accomplish pattern association of this type is variously referred to as the Widrow–Hoff learning rule (Widrow and Hoff, 1960) or the delta rule (Rumelhart and McClelland, 1986). This procedure uses the difference between a desired or 'target' activation and the obtained activation to guide learning in the system. In the present model, the target activation for the OL unit corresponding to a word's gender was set at 1, and the target activation for the other OL unit was set at 0.

The learning rule can be stated as:

$$w_{ij}(t) = w_{ij}(t-1) + re_i a_j \qquad (2)$$

where $w_{ij}(t)$ refers to the new weight created at trial t between OL u_i (u = unit) and IL u_j, $w_{ij}(t-1)$ refers to the prior weight for this connection, r represents a learning rate parameter, and e_i represents the difference, or 'error' for OL u_i, given by:

$$e_i = t_i - a$$

where t_i is the target activation on OL u_i and a_i is the obtained activation computed as in equation (1). In other words, the new weight equals the prior weight plus some fraction (the size of which is determined by the learning rate parameter) of the difference between the actual and desired weights.

An appropriate set of weights thus evolved gradually as the training procedure continued through several sweeps (or 'epochs') through the wordlist. The success of the network in learning the set of stimulus patterns was determined after each learning epoch by measuring the activations produced on the output units by each of the stimulus patterns. An item was thus considered accurately classified if the OL activation was greater on the appropriate gender node for its item class. The simulations of the model were performed on an Apple computer programmed in Applesoft BASIC.

III Computational experiments

1 Zero-state 'acquisition' of the stimulus set

Several experiments were performed to assess the ability of the network to learn the genders of the training-set of words under different conditions. In the first experiment, the model was initialized with its connection weights all set to zero and its learning rate set to 0.01. It was then cycled five times through the training set of words. In each cycle, for each word, the network first guessed the gender of the word, and then was presented with the word's correct gender designation. To the degree to which there was any discrepancy between the guess and the correct gender designation, the network revised its connection weights in accordance with the equation specified above. The current values of the connection weights were saved after each cycle. At the end of this training session, the model's weight-modification routine was disabled, and the connection weights that were previously saved were then utilized by the model to guess the gender of the study-set items after different degrees of learning. The overall accuracy of the model's choices after each training cycle are shown in Figure 2. After one pass through the training set, the network was able to correctly classify over 80% of both feminine and masculine words. By the fifth pass through the training set, accuracy had reached a level of 86.2% correct overall.

Thus, after five training cycles, the network was able to accurately classify all but 62 words out of the 450. However, in addition to accuracy, it was also possible to estimate the degree of 'confidence' the network displayed in each of its classifications by calculating the difference in activation values on the two OL nodes for each particular item. After five training cycles, the average difference

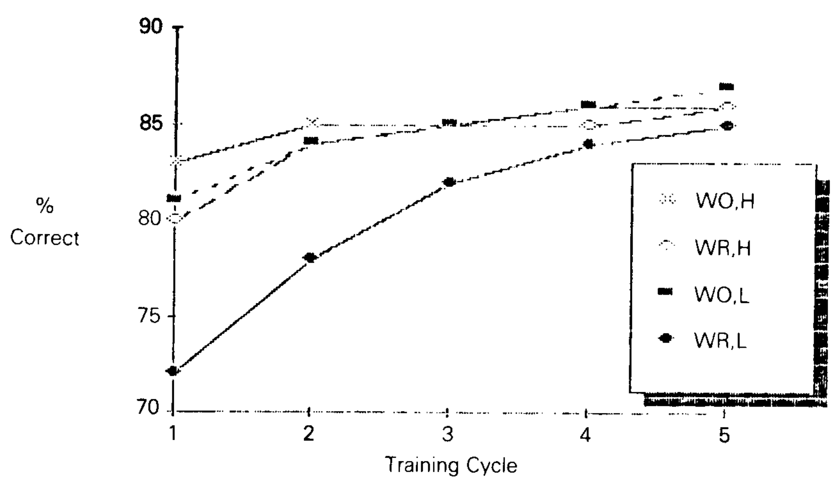

Figure 2 Gender identification accuracy for the network after different cycles of practice with the training set. WO = zero initial starting weights; WR = random initial starting weights; H = learning rate parameter set at 0.10; L = learning rate parameter set at 0.025.

Table 1 Classification 'confidence' for misclassified words

	>0.6	0.6–0.41	0.4–0.21	<0.21
1	conte	modèle	empire	oignon
2	automne	mer	verbe	hâte
3	cuiller	angle	prison	société
4	orange	image	nappe	auteur
5	montre	film	ennui	regle
6		mouchoir	liberté	chasse
7		coin	manque	bouche
8		forêt	spatule	flèche
9		lecteur	mur	branche
10		poêle	autorité	préfixe
11		mort	thé	esprit
12		acteur	saison	train
13		page	retour	fenêtre
14		part		lettre
15				bonté
16				tête
17				pomme
18				soir
19				toit
20				musée
21				couleur
22				saison
23				fait
24				santé
25				carotte
26				plaisir
27				est
28				classe
29				fois
30				glace

between activation levels on the two OL nodes was 0.682 for correctly classified items. The range of such confidence values for incorrectly classified nouns are displayed in Table 1. As shown in the table, only five items out of the 450, or about 1% of the total, were misclassified as confidently as the average confidence for correct items.

Most of the weights which evolved in this experiment were below an absolute value of 0.25, and less than 3% of the weights were higher than 0.50. Table 2 examines the distribution of weights that connected the last three letter positions to the masculine and feminine OL nodes, that were greater than an absolute magnitude of 0.24, and that existed after the fifth training cycle. As is apparent in the table, no single feature is critical to the model's categorization capabilities. Rather, the network has identified multiple features that can contribute relevant evidence to the gender identification task.

Table 2 Distribution of item connection strengths between gender nodes and input letters and positions

		Feminine			Masculine	
	Pos 6	Pos 7	Pos 8	Pos 6	Pos 7	Pos 8
a	0.28					
b	−0.45			0.67	0.34	
c				0.45		0.56
d					0.41	0.43
e			0.48		0.33	
f			0.46		0.34	
g		−0.55			0.79	
h		0.37		0.34		0.41
i	0.28	0.32				
j						0.25
k						
l				0.27		0.42
m		−0.26	0.28		0.51	
n						0.28
o		0.42				
p				0.30		
q				0.37		
r					0.32	
s		0.25		0.30		0.43
t				0.28		0.37
u		0.36				0.61
v				0.25	0.45	
w						
x	−0.31			0.40	0.35	
y				0.40		
z	−0.34			0.41		
è	0.33					
é	0.31					

2 *The influence of a pre-existing pattern of connectivity*

When the network started out with a zero-state in its pattern of connectivity, it quickly learned to classify the gender of words with a high degree of accuracy. However, it is reasonable to assume that a person learning gender associations comes to the task with some pre-existing pattern of connectivity. Indeed, if a person is trying to learn gender associations in a second language, we might assume that he or she typically has some well-established pattern of neural connectivity that is (at least, for example, in the case of the native English speaker) unrelated to the cognitive structures necessary to recognize the gender of nouns in the second language, and further, that will influence the way new information in the second language is neurally organized (cf. Vaid, 1983).

Therefore, a second experiment was conducted to examine the ability of the network to learn the genders of the stimulus set items when network connectivity was initialized by assigning a random value to each connection between the IL and the OL. An examination of the values of the interconnection weights produced in the first experiment indicated that most of them were close to zero, and few exceeded an absolute magnitude of 0.5. Thus, in this second experiment, the pre-assigned connection values ranged between negative 0.5 and positive 0.5 in strength, with a mean value of zero. All other aspects of the training procedure were identical to those described above.

In order to ensure that the results were stable, the network was trained with three different sets of random starting weights and the results were averaged across these three runs. Overall results from these experiments are compared across training cycles (as well as to the results of first experiment) in Figure 2. Network performance in the random-weight condition was slightly lower than in the zero-weight condition after cycle one, but thereafter performance for the two conditions converged. Thus, when all other conditions are equal, pre-existing weights only trivially interfere with the acquisition of gender associations by the model.

3 The influence of lowering the learning-rate parameter

Although in most domains adults appear to be much more capable of reasoning with abstract systems of formal rules than are children (Inhelder and Piaget, 1958), empirical studies (Johnson and Newport, 1989) have tended to confirm the common assumption that second language acquisition progresses more rapidly in children than it does in adults. Sokolik (1990) refers to this state of affairs as the 'Adult language-learner (ALL) paradox', and suggests that the relatively greater difficulty displayed by adult language-learners might reflect some age-related decrease in a connection learning-rate parameter, analogous to the one found in equation (2).

Thus, to approximate more closely the task confronting the adult learner of French, Experiments 1 and 2 above were repeated with the learning-rate parameter decreased from 0.1 to 0.025. The results from these subsequent studies are also displayed in Figure 3. When the network performs the learning task in the zero-state, the lower learning-rate has no appreciable effect. However, when the network performs the learning task with a random pre-existing set of connections and a low learning-rate (and thus is configured in a manner presumably analogous to the adult second-language learner), accuracy is initially much lower than in the other conditions, and it only begins to approximate the performance obtained in the other cases after several learning cycles.

4 Generalization of learning to new words

As noted in the introduction, pattern associators learn parameters of the statistical structure of a set of inputs. To the extent to which words of a particular gender share certain features, we would expect that the network model should also be able to generalize learning derived from one set of words to words it had never seen before. To examine directly whether the model was capable of generalizing the knowledge about gender assignment that it had acquired during training, the 150 (75 masculine and 75 feminine) stimulus items that were not utilized during the training cycles in Experiments 1–3 were submitted for classification, utilizing the cycle five weights obtained from each of the conditions described above.

The accuracy level achieved for these items are presented for each learning condition in Table 3. As can be seen in the table, the model was capable of predicting the gender of over 76% of new stimulus words when averaging across learning conditions. The existence of pre-existing connections did not appreciably influence the ability of the model to generalize gender associations to new words. However, when the model was trained under low learning-rate conditions, it was about 6% more efficient at classifying new words than when it was trained under higher learning-rate conditions.

Table 3 Proportion correct nonstudied words

	Fast learning	Slow learning
zero weights	0.733	0.787
random weights	0.731	0.793

IV General discussion

In the computational experiments described above, we have outlined a relatively simple model that is capable of learning the gender of a large set of French nouns. It accomplished this without relying on article or adjective agreement, without knowledge of noun meaning, and without being programmed with (or inferring) explicit morphological or phonological rules of gender formation. Rather it 'learned' that certain features (in this case, orthographic) of French nouns are correlated with particular genders. Based only on this information, it was able to classify at a high rate of reliability the gender of nouns it had never before encountered.

These studies provide evidence that gender can in principle be assigned during relatively low-level perceptual analysis without the application of explicit rules. Given that vocabulary items may be studied in relative isolation by the second language learner, it seems likely that the development of such pattern detection skills is also a normal component of the process by which the ability to assign gender to nouns is acquired. Inasmuch as article systems are a late learned feature in SLA (Master, 1987), pattern detection skills of this type might well provide a more reliable basis for gender assignment in the low-level second language learner than do contextual constraints. Indeed, given that French nouns may often appear with inherently ambiguous articles, like the plural form of the definite article, les, or the elided form in front of vowels, l', context may often provide an inadequate source of constraint even for the native speaker (Tucker *et al.*, 1968).

Further, the computational approach utilized here provides a naturalistic explanation of the common notion that language acquisition proceeds more slowly in adults (cf. Sokolik, in press). In brief, acquisition of gender was only slowed when the network was simultaneously hampered by a pre-existing set of weights that were unrelated to the task to be learned, and by a lower learning rate parameter – both conditions that might be expected to exist in the adult language learner. The simulations conducted here even led to a testable prediction about the adult second language learner's ability to generalize learning to new French nouns that was both nonintuitive and unexpected. More specifically, the model's behaviour predicts that once the second language learner has acquired a native speaker's ability to identify the gender of familiar nouns, he or she should also demonstrate a superior ability to correctly identify the gender of unfamiliar nouns. This ability to generate such nonintuitive predictions attests to the strength of formal modelling approaches in the study of second language acquisition.

Finally, pattern detection mechanisms of the general form described here might exist for other linguistic classification functions. Convergent evidence for this possibility is provided by studies of 'implicit learning' (for a review see Reber, 1989). Such studies have illustrated that knowledge about many aspects of the regularities inherent in a stimulus environment can be acquired independently of efforts to infer the rules that govern those regularities. For example, in the domain of grammaticality judgements, Reber (1976) found that subjects were not aided in deciding whether a given letter string was produced by a finite-state grammar when they were instructed to infer the rule structure of that grammar, nor when they were explicitly provided with that rule structure by the experimenter. Instead, optimal performance was achieved simply by showing subjects many examples of strings that were produced by the grammar. Further, the knowledge acquired in such experiments can be utilized independently of the ability to articulate the principled factors that are governing behaviour (e.g., Lewicki *et al.*, 1988). Performance by subjects in these situations then closely resembles that inherent in making assignments of noun gender. Indeed, Tucker *et al.* similarly suggest that 'the French speaker is sensitive, although apparently at some level far below awareness, to these recurring patterns of regularities in his language' (1968: 315). Thus, such correspondence in behaviour across diverse tasks implies that pattern detection mechanisms of the form described herein might well contribute to more sophisticated forms of linguistic processing as well.

Acknowledgements

A portion of this research was done while the authors were at the Centre Paul Broca, INSERM U97, Paris, France. Some of the results were presented at the *Second Language Research Forum*, UCLA, Los Angeles, CA, February, 1989.

VI Appendices

Appendix A Feminine gender words

abbaye	banque	cave
abeille	barbe	caverne
actrice	bergère	centaine
action	bière	cerise
addition	boite	chaise
adresse	bonté	chance
affaire	bouche	chapelle
affiche	bourse	charge
affinité	boutique	charité
aise	branche	chasse
alliance	brochure	cheminée
année	caméra	chemise
annonce	campagne	chimie
armoire	capacité	chose
assiette	capitale	chute
auberge	carotte	classe
autorité	carrière	clef
aventure	carte	coiffure
avenue	case	colline
banane	cause	comédie

compagne
conquète
copie
côte
couleur
cour
craie
cravate
création
crème
crise
cuiller
cuisine
dame
date
décision
défaite
demande
destinée
dette
dizaine
douzaine
droite
eau
école
écriture
édition
église
élection
élégance
enfance
énigme
entrée
épée
épithète
époque
ère
erreur
essence
étendue
étude
excuse
fabrique
faim

famille
fatigue
faute
fée
fenêtre
ferme
fête
feuille
figure
fin
finesse
flèche
fleur
foi
fois
fonderie
fondue
force
forêt
forme
fortune
foule
fumée
galette
gare
glace
gloire
gravure
guerre
habitude
haleine
harpe
haste
hauteur
herbe
heure
histoire
honte
huile
idée
identité
île
image
initiale

insomnie
invasion
issue
jambe
jeunesse
journée
laitue
langue
lanterne
latitude
lavande
lampe
lecture
lettre
liberté
licence
ligne
liste
loi
lumière
lune
machine
main
maison
manière
marche
marée
marque
marquise
matière
médaille
médicine
mer
miette
minute
moitié
monnaie
montagne
montre
mort
musique
nappe
nation
nature

noblesse
note
notice
nuit
opinion
orange
oreille
origine
page
paille
parole
part
partie
passion
patience
pause
pêche
peine
pelouse
pendule
période
personne
perte
peur
phrase
pièce
pierre
pipe
place
planète
plaque
pluie
poche
poêle
police
pomme
porte
portion
poubelle
poule
présence
prière
prise
prison
promesse
province
qualité
question
queue
radio
raison
réalité
recette
régence
région
règle
religon
réponse
réunion
revue
rivière
robe
rose
roue
rouille
roulette
route
rue
ruine
saison
salade
salle
santé
sauce
section
semaine
série
société
soie
soif
soirée
spatule
terre
tête
théière
théorie
thèse
toiture
tomate
trace
tuile
valeur
vallée
veille
vérité
viande
victime
victoire
vie
ville
visite
vitesse
vitrine
vogue
voie
voiture
volonté
vue

Appendix B Masculine gender words

abîme
accident
accord
acte
acteur
adieu
affront
âge
air
allié
amateur
ancêtre
angle
animal
annuaire

apéritif
appareil
appétit
arbre
arc
argent
aspect
asphalte
assassin
auteur
autobus
automne
avion
avis
avocat
bagnol
bain
bal
barbier
bateau
beurre
bifteck
bijou
biscuit
bois
bonheur
bouchon
bouquin
bout
bruit
buffet
bureau
cabinet
cadeau
café
canal
canton
cargo
cas
cendre
centime
champ
change
chapeau
chat
château
chef
chèque
cheval
chien
choix
chou
ciel
circuit
climat
coin
collège
complet
compte
congé
congrès
conte
coup
courage
cours
couteau
crayon
cri
décor
degré
déjeuner
déluge
départ
député
dessert
détail
détour
dieu
discours
disque
document
doigt
dôme
drapeau
droit
écho
écran
écrit
écrivain
effet
effort
empire
ennemi
ennui
ensemble
esprit
est
estomac
étang
état
été
excés
exercise
fait
fer
feu
film
fleuve
fond
fou
franc
fromage
futur
garage
garde
gaz
genre
geste
goût
grain
grenier
grillon
hasard
hiver
honneur
horizon
hôtel
immeuble
incident
individu
instant
invité

jardin
jouet
journal
juge
jus
lâcher
lait
lecteur
lieu
linge
liquide
litre
loup
magasin
maître
malheur
manque
marché
matin
médecin
menu
message
métal
midi
milieu
million
modèle
mois
moment
monstre
mont
morceau
moteur
mouchoir
moulin
mur
musée
mystère
nom
nombre
nord
objet
offre
oignon
opéra
ordre
ouest
palais
pantalon
papier
paquet
parc
parent
pas
passé
patron
paysage
permis
peuple
pistolet
placard
plaisir
plat
plateau
poème
poisson
poivre
pont
port
portrait
potage
poulet
pouvoir
préfixe
problème
procédé
projet
quart
quartier
rameau
reflet
regard
régime
rampart
repas
repos
retard
retour
rêve
rhume
rideau
rien
rival
rocher
rôle
rosbif
rôti
sac
salon
sang
savant
séjour
semestre
sénat
service
siècle
signal
soir
soleil
sommeil
soulier
souper
soupir
stylo
succès
sucre
suicide
sujet
symbole
tableau
tabouret
tapis
teint
temps
terme
texte
thé
théâtre
tiroir
titre
toit
tourisme

train	verbe	reste
triomph	vin	plan
trône	visage	nuage
trou	vol	lit
type	voyage	excès
usage	volume	remède
velours	timbre	vignoble
vent	silence	

VII References

Andersen, R. 1984: What's gender good for anyway? In Andersen, R., editor, *Second languages: a cross-linguistic perspective.* Rowley, MA: Newbury House, 77–100.

Anderson, J.A. 1983: Cognitive and psychological computation with neural models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC–13*, 799–815.

Clark, A. 1989: *Microcognition: philosophy, cognitive science, and parallel distributed processing.* Cambridge, MA: Bradford Books.

Clark, E.V. 1985: The acquisition of Romance, with special reference to French. In Slobin, D.I., editor, *The crosslinguistic study of language acquisition*, Volume 1. Hillsdale, NJ: Lawrence Erlbaum & Associates, 687–782.

Hinton, G.E. and **Anderson, J.A.** editors, 1981: *Parallel models of associative memory. Hillsdale, NJ: Lawrence Erlbaum & Associates.*

Inhelder, B. and **Piaget, J.** 1958: *The growth of logical thinking from childhood to adolescence; an essay on the construction of formal operational structures.* New York: Basic Books.

Jacobs, B. 1988: Neurobiological differentiation of primary and secondary language acquisition. *Studies in Second Language Acquisition* 10: 303–38.

Johnson, J.S. and **Newport, E.L.** 1989: Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99.

Kohonen, T. 1984: *Self-organization and associative memory.* Berlin: Springer-Verlag.

Lewicki, P., Hill, T. and **Bizot, E.** 1988: Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology* 20: 24–37.

Master, P.A. 1987: A crosslinguistic interlanguage analysis of the acquisition of the English article system. Ph.D. Dissertation, UCLA.

McClelland, J.L. and **Rumelhart, D.E.** 1986: *Parallel Distributed Processing: explorations in the Microstructures of Cognition*, Volume 2: *Psychological and biological models.* Cambridge, MA: MIT Press.

Pinker, S. and **Prince, A.** 1988: On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. In

Pinker, S. and Mehler, J., editors, *Connections and symbols*. Cambridge, MA: MIT Press, 73–194.

Reber, A.S. 1976: Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory* 2, 88–94.

— 1989: Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General* 118, 219–35.

Rumelhart, D.E. and **McClelland, J.L.** 1986: *Parallel Distributed Processing*, Explorations in the Microstructures of Cognition Volume 1: *Foundations*. Cambridge, MA: MIT Press.

Sokolik, M.E. 1990: Learning without rules: PDP and a resolution of the adult language learning paradox. *TESOL Quarterly* 24, 685–96.

Surridge, M.E. 1986: Genre grammatical et derivation lexicale en français. *Canadian Journal of Linguistics* 31, 267–83.

Stevens, F. 1984: *Strategies for second-language acquisition*. Montreal: Eden Press.

Tucker, G.R., Lambert, W.E., Rigault, A. and **Segalowitz, N.** 1968: A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior* 7, 312–16.

Vaid, J. 1983: Bilingualism and brain lateralization. In Segalowitz, S., editor, *Language functions and brain organization*. New York: Academic Press.

Valdman, A. 1969: *An introduction to French phonology and morphology*. Rowley, MA: Newbury House.

Widrow, G. and **Hoff, M.E.** 1960: Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record* Part 4, 96–106.