Grammatical gender assignment in French: dispelling the native speaker myth

DALILA AYOUN

University of Arizona

ABSTRACT

This study highlights the complexity of French grammatical gender as a lexical property at the interface of morpho-phonology and the lexicon. French native speakers (n = 168) completed a gender assignment task with written stimuli illustrating common versus uncommon nouns, vowel-initial versus consonant-initial nouns, compounds and grammatical homonyms; they also indicated the strategies they used to assign a gender to stimuli. The findings showed strong lexical and gender effects suggesting that grammatical gender must be acquired for individual lexical items as morpho-phonological cues alone are unreliable and vary greatly.

INTRODUCTION

It has long been assumed that native speakers (NS) of a language with an opaque gender system such as French can accurately attribute the appropriate gender(s) to existing nouns as well as pseudo-nouns (Tucker, Lambert, Rigault and Segalowitz, 1968; Tucker, Lambert and Rigault, 1977) by relying on morpho-phonological endings and their intuition or "through experience, one just knows it, that's all" (Tucker, 1968: 6). Empirical evidence shows that grammatical gender knowledge emerges shortly after the first year (Höhle, Weissenborn, Kiefer, Schulz and Schmitz, 2004; Shi and Melançon, 2010), and that infants use the gender features of determiners to assign gender to pseudo-nouns (Cyr and Shi, 2013), while 30-month-old toddlers are faster during gender-matched trials than gender-unmatched trials with real and pseudo-words (Melançon and Shi, 2015). By the age of 3, children rarely make errors in attributing gender to nouns (Clark, 1985). However, French displays idiosyncracies, incongruences and asymmetries in both gender attribution and gender agreement (e.g., Hulk and Tellier, 1998, 2000; Schafroth, 2003; Wagner and Pinchon, 1991), which, according to anecdotal evidence, may present difficulties even for NSs.

To the best of my knowledge, the only empirical data supporting this claim were collected for a second language (L2) acquisition study of grammatical gender and agreement by Anglophone learners (Ayoun, 2007). The data from the French NS control group suggested that their judgments may not be as reliable as they

are assumed to be, since a strong lexical effect was found, particularly with nouns exhibiting fluctuating gender. For instance, the results of a written gender assignment task (GAT) revealed significant differences between adult (n=20) and teenage (n=42) participants in that adults agreed more often on the gender of stimuli than teenagers. Thus, all adults rated 21 out of 50 nouns (54%) as feminine, but teenagers agreed 100% of the time on only one noun (*cible* 'target'). Among nouns with very low accuracy percentages were *idole* 'idol' (42.9% for teenagers), *oasis* 'oasis' (50.0% for adults, 16.7% for teenagers) and *primeur* 'scoop' (42.9% for adults, 2.4% for teenagers). Moreover, participants appeared to rely on questionable cues from determiners and spelling word endings (Ayoun, 2009).[1]

   The present study may be the first since Tucker et al. (1968, 1977) to focus on French NSs (n=168) to investigate their performance on: a) common and less common nouns; b) dual gender nouns (i.e., grammatical homonyms); c) compounds whose gender almost always differs from the gender of the second lexical item (or both). An important difference is that Tucker et al. (1968, 1977) used oral stimuli whereas the present study employs written stimuli. Although the gender of the lexical items remains the same, auditory stimuli may be processed differently than written stimuli. The main objectives of the present study are to determine: a) how reliable and consistent French NSs are at assigning the appropriate gender to common, uncommon nouns and compounds in a written gender assignment task; b) what underlying strategies may be used in doing so. It will be hypothesized that purely associative mechanisms are not sufficient and that gender must be acquired for each lexical item as morpho-phonological cues alone are unreliable. Lexical knowledge also comes into play in the case of idiosyncracies (i.e., fluctuating gender, grammatical homonyns/homophones, epicenes, compounds). Following an overview of the French gender system, selected studies will be reviewed before presenting the methodology, a discussion of the findings, and a few conclusions regarding both native speakers and second language (L2) learners.

GENDER ASSIGNMENT IN FRENCH

French is described as having a formal gender system with a weak semantic component because it exhibits both inherent lexical (or semantic) gender and grammatical gender, while other languages such as English have a system in which gender is not grammatically marked, but semantically motivated (Corbett, 1991, 2003; Séguin, 1969).[2] In such a system, "semantic factors are sufficient on their own to account for assignment" (Corbett, 1991: 8): nouns referring to male human beings are masculine, while nouns referring to female human beings are feminine.

---

[1] van der Linden and Hulk (2009) also show that NSs may have difficulties.
[2] The fully productive inflectional gender system of Old English has been reduced to a few gender markers such as the pronouns 'it/its' (neutral), 'he/him' (masculine) and 'she/her' (feminine), although 'they' as a singular, gender neutral form tends to be used instead of 'he/she' or a few nouns such as 'actor/actress' or 'waiter/waitress' (Kastovsky, 2000).

The gender that is assigned to all inanimate nouns in French is generally referred to as arbitrary because "grammatical gender for inanimate nouns has no conceptual basis, and does not go through a process of feature marking in mapping the conceptual notion onto a linguistic code" (Deutsch and Dank, 2009: 116). In addition to a semantic (i.e., biological) gender and an arbitrary gender, French displays: a) a morphological gender at the word level from the same lexical entry (e.g., *Italien* 'Italian-MSC', *Italienne* 'Italian-FEM', *pharmacien* 'pharmacist-MSC', *pharmacienne* 'pharmacist-FEM'); b) a relational syntactic gender between phrasal constituents used in syntactic concord as in *la jolie fleur violette* 'the-FEM beautiful-FEM purple-FEM flower' where the determiner and adjective must agree in gender (and number) with the noun; c) referential gender for *épicènes*, nouns that are either masculine or feminine depending on their referent (e.g*., un/une journaliste* 'a– MSC/FEM journalist').

According to Séguin (1969), semantically motivated gender categories account for only 10.5% of all French nouns, which means that 89.5% of nouns display a grammatical gender. Séguin also points out that there is an uneven distribution of masculine and feminine nouns since 58.6% of nouns are masculine, creating a potential bias toward the masculine, as discussed below. The perennial question is whether grammatical gender is assigned on a principled basis or is completely arbitrary (see e.g*.,* Carroll, 2005; Maratsos and Chalkey, 1980; Müller, 1990).

How is grammatical gender assigned?

A variety of gender assignment rules were proposed that were first based on semantics and spelling (Bidot, 1925), then on phonological endings accounting for roughly 85% of frequently occurring nouns (Mel'čuk, 1958), as well as a combination of semantic and phonological rules (Tucker et al., 1968, 1977). In the comprehensive system Tucker et al. (1977) proposed based on the lexicographical analyses of 26,725 nouns from the *Petit Larousse*, most rules are phonological, some are semantic and morphological, while different types of rules may overlap or conflict, and exceptions abound (e.g., Surridge 1993). Indeed, the French system "allows more exceptions than do other systems" (Corbett, 1991: 61).[3] Tucker et al. (1977) contend that spelling also plays a role; for instance, *-é* and *-ée* yield both masculine (e.g., *lycée* 'high school', *musée* 'museum', *comité* 'committee') and feminine nouns (e.g., *mosquée* 'mosque', *journée* 'day'). However, their own generalizations are plagued with counter-examples that they dismiss as "inappropriate" (*ibid*: 61), referring for instance to the fact that most nouns ending in *-age* are masculine, with the notable exceptions of *page* 'page', *image* 'image', *cage* 'cage', among others.

---

[3] For more detailed critical accounts of Tucker et al. (1977) see Antes (1993), Beckett (2010), Carroll (1989) or Warden (1997); the consensus appears to be that the rules they offer are an *ad hoc* summary of statistical regularities and that the phonological endings are "only weakly probabilistic" (Cyr and Shi, 2013).

Another way to use morphology is to draw parallels between suffixation and gender assignment (e.g., Gervais, 1993; Surridge, 1986, 1993), as in the following examples (taken from Surridge, 1986: 274):

(1)   a. Nouns derived from adjectives with the following suffixes are feminine:
        -eur, -ie, -ité, -icité, -esse, -itude, -étude, -ance, -ence, -ise
      b. Nouns derived from nouns with the following suffixes are feminine: *-elle, -ette, -elette, -iole, -ule, -aie, -eraie, -aine*
      c. Nouns derived from other nouns with the following suffixes are masculine: *-eau, -ot, -on, -eron, -in, -et, -illon, -icule*
      d. Nouns derived from the suffixe *–isme* are masculine

Relying on suffixation requires a solid knowledge of morphology since only suffixes provide deterministic cues to gender, while similar endings do not (MacWhinney, 2000). Thus, *conversation* 'conversation' is feminine because it ends with the suffix **-tion,** whereas *bastion* 'bastion' is masculine, because it is not composed of *bas-+-tion* (Presson, MacWhinney and Tokowicz, 2014).

Semantic rules appear to be used when phonological and morphological accounts fail. Thus, the semantic concept of 'days of the week' and 'seasons' is used to explain why all the words subsumed under these categories are masculine (Surridge, 1993). Beckett (2010: 57) identifies five lexical fields: birds, fish, "other members of the animal kingdom", "plant kingdom, limited to two areas – woody plants (trees, shrubs, vines) and fruits", and human beings. Although Beckett pursues the "potential relationship between phonology and gender assignment [. . .] in relation to word-final syllable structure and morphology" (*ibid*: 57), it is argued that semantic features related to size, form or shape are helpful in accounting for gender assignment. In addition to the caveat that only 8,000 words were used, it is difficult to see how an account limited to only two categories of nouns can be useful beyond a semantic analysis of these very specific nouns.[4]

One could argue that the masculine acts as the default gender because 58.6% of nouns are masculine, and because it is attributed to several categories such as: a) nominalized verbs (*boire* 'drinking', *déjeuner* 'lunch', *lancer* 'throw', *pouvoir* 'power', *rire* 'laughter', *savoir* 'knowledge', *toucher* 'touch', *vouloir* 'will', *savoir-faire* 'know how'*). However, nouns derived from verbs are feminine (*entrée* 'entrance', *allée* 'alley' *arrivée* 'arrival', *fumée* 'smoke', *mêlée* 'fray'); b) nominalized adjectives and participles are masculine (*chaud* 'hot', *froid* 'cold', *clair* 'clear', *bon* 'good', *mauvais* 'bad'), as are generally words borrowed from other languages ('judo', 'spoutnik', 'gang'), although Desrochers (1986) argues that gender assignment for loan words can be either semantic or phonological.[5] Moreover, and in a similar way to the semantic analysis of very specific nouns, because such rules of thumb concern a very

---

[4] Beckett (2010) gets into complex semantic distinctions in an effort to establish a systematic gender assignment such as the way organic matter is perceived or the various features attributed to animated beings such as eagles or their mode of existence.

[5] Enger (2009: 1286) warns us against the danger of using "notions such as 'default' and 'marked' as a shorthand label for factors we do not understand", addding that "Kilarski

small subset of the data, they are not very useful; c) English borrowings tend to be masculine, but out of a corpus of 703 nouns drawn from popular music magazines, 51.49% were masculine, 12.09% feminine, while 36.2% were ambiguous (Guilford, 1994, 1999; Pergnier, 1989 reports similar findings). Gender appears to often be assigned by semantic analogy (e.g., *la house*-FEM for 'house music' by analogy with *la maison*-FEM);[6] d) Foreign nouns code-switched in French also appear to be overwhelmingly assigned the masculine gender (Violin-Wigent, 2006).

Furthermore, gender conflict resolution always favors the masculine in standard usage as shown in (2):

(2)   a. *Un cousin et une cousine amusants*
          a–MSC cousin–MSC and a–FEM cousin–FEM fun–MSC–PL
       b. **Un cousin et une cousine amusantes*
          a–MSC cousin–MSC and a–FEM cousin–FEM fun–FEM–PL
          'Fun cousin and cousin'

This creates some semantic ambiguity (Irmen and Kurovskaja, 2010) as in (2a) where it is unclear whether both people are fun or not. It is generally resolved by assuming that a masculine interpretation is favored over a feminine interpretation in French (e.g., Gygax, Sarrasin, Lévy, Sato and Gabriel, 2013; Gygax and Gabriel, 2008), as in other languages such as German or Norwegian (e.g., Gabriel, 2008; Irmen and Schumann, 2011; cited in Gygax et al. 2013).

Tables 1 and 2 display oral data compiled from Tucker et al. (1977) to highlight the predictive value of consonantal and vocalic final phones. Both tables are organized by the decreasing predictive value of the phones (% column).[7] Also given is the number of tokens for each phone, an important piece of information for if a phone has a high predictive value, but few tokens, its predictive value is reduced; conversely, a low predictive value is compounded by a large number of tokens. The 'spelling' column indicates all the possible forms for each phone.

Table 1 shows that only 3 out of 18 phones have a predictability value of 90% or above (a reliability criterion also used by Lyster, 2006), while 4 of them hover around 50% to 58%, including phones with large number of tokens such as [l] and [t], as well as numerous different spellings.

The data displayed in Table 2 reveal much better predictability values for vocalic final phones, although the number of tokens is smaller, particularly for [œ̃] (only 17 tokens) and [ø] (only 189 tokens compared to 1963 tokens for [ã] for instance).[8]

(2001) even calls the concept of a default in gender assignment 'a dustbin category of no explanatory value'". See also Roché (1992) or Ayres-Bennett and Carruthers (2001).

[6] Aside from semantic analogy, gender assignment to borrowings may also follow cross-linguistic analogy, morpho–phonological analogy and hyperonymy (Anastassiadis-Syméonidis, 2005; Anastassiadis-Syméonidis and Nikolaou, 2011) (I am grateful to an anonymous reviewer for these references).

[7] These tables and the description of idiosyncracies are partially adapted from the tables in Ayoun (2007, 2010).

[8] Indeed, the [œ̃]/[ɛ̃] distinction is increasingly disappearing in some regional variations of French, notably in the Paris area.

Table 1. *Consonantal final phones, tokens and spellings*

phone	spelling	#tokens	M	F	%
[ʒ]	–j, –ge, –ges	1453	1368	85	94%
[m]	–m, –me, –mes	1406	1292	114	92%
[z]	–se, –ze	612	61	551	90%
[r]	c+re, +res, v+[r]+c	512	417	95	81%
[f]	–f, –fs, –fe, –fes, –phe, –phes	131	101	30	77%
[g]	–g, –gs, –c, –gue, –gues	235	172	63	73%
[v]	–v, –ve, –ves	143	45	98	69%
[j]	–v+il, ille, illes	352	114	238	68%
[n]	–n, –ne, –nes	1135	358	777	68%
[d]	–d, –de, –des	668	227	441	66%
[ʃ]	–ch, –sh, –che, –ches	290	99	191	66%
[b]	–b, –be, –bes, – bbe, –bbes	129	84	45	65%
[ɲ]	–gne, –gnes	69	27	42	61%
[s]	–s, –ss, –x, –ce, –se, –xe, –ces	1379	531	848	61%
[l]	–l, –ls, –le, –les, –lle, –lles	1126	474	652	58%
[k]	–c, –cs, –ch, –chs, –ck, cks, –q, –que, –ques	609	333	276	55%
[p]	–p, –pe, –pes	214	104	110	51%
[t]	–t, –te, –tte, –the, –tes, –ttes, –thes	2269	1162	1107	51%

Table 2. *Vocalic final phones by tokens*

phone	orthography	#tokens	M	F	%
[œ̃]	–um, –un, uns, –unt	17	17	0	100%
[ã]	–an, –anc, –and, –anf, –ang, –aon, –amp, –ans, –ancs, –amps, –ant, –end, –eng, –ens, –ends, ems, –empt, –ent, –ents, –ants	1963	1949	14	99%
[ɛ̃]	–aim, –ym, –én, –ien, –en, –ain, –ein, –in, aing, –oing, –eing, –éens, –iens, –ains, –eins, ins, –inct, –ingt, –ient, –aint, –eint, –int	938	929	9	99%
[ø]	–oeud, –eue, oeufs, –eut, –eu, –eux	189	184	5	99%
[o]	–o, –oc, –op, –os, –ots, –ot, –aud, –aut, –ault, au, –aux, –eaux, –aulx	865	841	24	97%
[y]	–u, –ul, –us, –uts, –ut, –ux, –ue, –ues	201	195	6	95%
[u]	–ou, –ouc, –oul, –oo, –oup, –ous, –out, –ouls, oux, –oue, –oues	171	150	21	88%
[wa]	–oi, –oids, –ois, –oigt, –oit, –oix, –oie, –oies, oye	179	153	26	85%
[a]	–a, –ac, –ap, –ats, –as, –at	791	648	143	82%
[i]	–ic, –id, –il, –is, –it, –ix, –i, –y, –ys, –ie, –ies, ye	2337	575	1762	75%
[õ]	–on, –om, –on+c, –om+c	2668	794	1874	70%
[e]	–é, –és, –ée, –ées, –er, –ers, –ez, –ai, –ais, –ait, aits, –aix, –aie, –aies, –ay, –et, –êt, –ès, –ect, ects, –ey, –egs	3416	1962	1454	57%

Moreover, [e] with the largest number of tokens (3416) has the lowest predictability value (57%). With the exception of [i], [e] and [ɔ̃], vocalic phones are masculine, whereas consonantal phones are more evenly split between masculine and feminine.

Lyster (2006: 72) also attempts to "determine the extent to which noun endings in French are reliable predictors of grammatical gender" with a written corpus. Using 90% as a reliability criterion, the analysis of the endings of 9,991 singular inanimate nouns selected from *Le Robert Junior Illustré* showed that only 6 out of 29 final phonemes (18% of the corpus) met that criterion (*ibid*: Table 1, p. 75). Noun endings were operationalized as spelling representations of rhymes.[9] Lyster argues that 81% of all feminine nouns and 80% of all masculine nouns in his corpus are rule-governed because their endings systematically predict their gender. However, the author is using a pre-selected corpus of nouns already classified as exhibiting typically masculine or feminine endings, and yet, only 81% of feminine nouns and 80% of masculine nouns within that corpus are indeed feminine or masculine. Moreover, dual gender and compound nouns were excluded. As previously noted, "arguably, nearly 400 spelling endings are too numerous to be of much use to L2 learners" (Tucker et al., 1977: 86).[10]

It thus may be best to accept that grammatical gender assignment in French is based on a mix of morpho-phonological and lexical rules in addition to spelling and semantics, rather than insist on "crazy rules" (Enger, 2009) that look like "post-factum rationalizations" (Comrie, 1999).

Idiosyncracies in gender

Although most nouns are masculine or feminine, grammatical homonyms are both, while the gender of a few other nouns fluctuates. Morevoer, although nouns referring to male human beings are masculine (e.g., *neveu* 'nephew'), while nouns referring to female human beings are feminine (e.g., *nièce* 'niece'), that is not always true of all animated beings. Thus, *abeille* 'bee', *baleine* 'whale', *mouche* 'fly', among many others, only have feminine forms and *mâle* 'male' is added to refer to the male species as in *une abeille mâle*, while *phoque* 'seal' or *singe* 'monkey', for instance, only have masculine forms, and *femelle* 'female' is added to refer to the female species as in *un singe femelle.*

Grammatical homonyms

Grammatical homonyms are pairs of nouns yielding a masculine form and a feminine form with identical spellings and pronunciation, but different meanings as exemplified in (3):

---

[9] Nucleus alone for a vowel, or nucleus plus a coda for a consonant plus a vowel.

[10] Séguin (1969) found 570 endings in a written corpus out of 18, 571 masculine words, and 128 feminine endings out of 13, 211 feminine words.

(3)   *un livre* 'a book'-MSC, *une livre* 'a pound'-FEM
      un moule 'a mold'-MSC, *une moule* 'a mussel'-FEM
      un manche 'a handle'-MSC, *une manche* 'a sleeve'-FEM
      le mémoire 'memoir, thesis'-MSC, *la mémoire* 'memory'-FEM

The gender fluctuation of some homonyms distinguishes between animate and inanimate nouns as in (4):

(4)   *aide* 'aid'-MSC, *aide* 'help'-FEM
      manœuvre 'worker'-MSC, *manœuvre* 'manoeuvre'-FEM
      enseigne 'ensign'-MSC, *enseigne* 'sign'-FEM
      secrétaire 'secretary'-FEM/MSC, *secrétaire* 'desk'-MSC
      page 'page'-FEM, *page* 'page'-MSC

Although the context disambiguates between their two possible meanings, these homonyms are clearly an exception to gender assignment based on phonological endings. There are about 50 grammatical homonyms (L'Huillier, 1999; Price, 2008).

Homophones

Homophones – words with different spellings, but identical pronunciation – form another category as exemplified in (5):

(5) [sɛl]   *sel* 'salt'-MSC, *selle* 'saddle'-FEM
    [fwa]   *foie* 'liver'-MSC, *fois* 'time'-FEM, *foi* 'faith'-FEM
    [ru]   *roux* 'redhead'-MSC, *roue* 'wheel'-FEM
    [rɛn]   *renne* 'reindeer'-MSC, *reine* 'queen-FEM
    [po]   *pot* 'jar'-MSC, *peau* 'skin'-FEM

Although the context eliminates any ambiguity, homophones are also a counter-example to gender assignment rules based on phonology. They are presumably stored with a single phonological representation with two different meanings (Gottlob, Goldinger, Stone and Van Orden, 1999). The interesting questions of how they may be stored lexically, and how they may be accessed, have been addressed by a few studies reviewed below (e.g., Spinelli and Alario, 2002).

Épicènes

Épicène nouns may be used with either grammatical gender depending on their referent, without a change in meaning or form:

(6)   *un/e* 'a-MSC-FEM' *artiste* 'artist', *juge* 'judge', *propriétaire* 'owner', *camarade* 'friend', *pensionnaire* 'boarder', *partenaire* 'partner', *stagiaire* 'trainee'

However, although *personne* 'person' is not an *épicène* per se, it could be argued that it is an exception in that it is always feminine regardless of its referent as in (7):

(7)   a. *Ma sœur/mon frère est une personne intelligente/*intelligent*
         my-FEM sister/my-MSC brother is a-FEM person intelligent-FEM/*MSC
         'My sister/my brother is an intelligent person'

b. *Sa tante et ses oncles sont des personnes charmantes/*charmants.*
his-FEM aunt-FEM and his uncles-MSC are det-PL persons-FEM
charming-FEM/*MSC
'His/her aunt and his/her uncles are charming people'

Other nouns with animate referents (but not *épicènes*) are only feminine or masculine regardless of the gender of the referent: *victime* 'victim', *sentinelle* 'sentry', *vedette* 'star', *connaissance* 'acquaintance' or *doublure* 'double, stand-in' are always feminine, while *génie* 'genius', *mannequin* 'model', *bébé* 'baby', *ange* 'angel', *témoin* 'witness', *ascendant* 'ascendant' are always masculine. Moreover, some titles such as *altesse* 'prince', *majesté* 'majesty', *Sainteté* 'Holiness' are feminine although their referents may be masculine as well as feminine (except for *Sainteté*).

Gender fluctuation with number

The gender of a few nouns – *amour* 'love', *gens* 'people', *délice* 'delight', *orgue* 'organ' – fluctuates with number. The first two are very common and quite interesting, as shown in (8):

(8)   a. *Les vieilles/*vieux gens sont méchants/*méchantes*
         the old-FEM/MSC people-PL are mean–MSC/FEM
         'Old people are mean'
      b. *Ces *belles/beaux jeunes gens sont méchants/*méchantes*
         these-FEM-MSC young-*FEM/MSC people are mean-MSC/*FEM
         'These young people are mean'
      c. *C'est l'histoire d'un/*une bel/*belle amour.*
         it is the story of a-MSC/*FEM beautiful-MSC/*beautiful-FEM love
         'It's a beautiful love story'
      d. *C'est l'histoire de *beaux/belles amours.*
         it is the story of det-PL *beautiful- MSC/beautiful-FEM love
         'These are beautiful love stories'

Gens 'people' is particularly idiosyncratic in that as a plural noun with either male and/or female referents, it triggers feminine agreement with preposed adjectives, but masculine agreement with postposed adjectives. Moreover, the phrase *jeunes gens* 'young people'– in which it appears that the adjective has been nominalized – is always masculine (the referents may be both feminine and masculine or only masculine, but not all feminine).

Compounds

Hawkins and Towell (2010: 17–19) identify six types of compounds to which gender is added here: a) adjective + noun compounds take the gender of the head noun (*rond-point* 'roundabout'-MSC), but there are exceptions (*rouge-gorge* 'robin'-MSC); b) in noun + noun compounds, the gender is based on the head noun (*mot-clé* 'a key word'-FEM); c) adverb + noun compounds have the same gender as the

head noun (*arrière-pensée* 'after thought'-FEM) so may be masculine or feminine, but with exceptions (*guerre* 'war'-FEM, *après-guerre* 'period after the war'-MSC); d) noun + prepositional phrase compounds take the gender of the first noun so may be masculine (*chef d'œuvre* 'masterpiece'-MSC, but *œuvre*-FEM) or feminine (*main d'œuvre* 'work force'-FEM); e) verb + noun compounds are usually masculine (*allume-cigarette* 'cigarette lighter'-MSC) regardless of the gender of the noun (*cigarette*-FEM); f) verb phrase compounds (*qu'en-dira t-on* 'what people may say'-MSC) are always masculine.

Thus, most types of compounds (particularly deverbal compounds) are masculine regardless of the gender of their lexemes as individual items as in (9):

(9) *croûte* 'crust'-FEM, *casse-croûte* 'snack'-MSC
    glace 'glass'-FEM, *essuie-glace* 'windshield wiper'-MSC
    faute 'error'-FEM, *sans-faute* 'perfect score'-MSC
    tête 'head'-FEM, *tête-à-tête* 'face-to-face'-MSC

A few feminine compounds are listed in (10):

(10) *fenêtre* 'window'-FEM, *porte-fenêtre* 'French window'-FEM
    temps 'time'-MSC, *mi-temps*-FEM 'half-time' (but also 'part-time'-MSC)
    part 'share'-FEM, *quote-part* 'share'-FEM
    face 'face'-FEM, *volte-face* 'flip-flop'-FEM
    garde 'watch'-FEM, *garde-à-vue* 'observation'-FEM

Both masculine and feminine compounds illustrate the fact that morpho-phonological endings are not reliable cues to gender because the same word can take on the opposite gender once it becomes part of a compound, although there are no morphological changes to its ending.[11]

To sum up, French exhibits a formal gender system in which 89.5% of nouns display an arbitrary gender. Morpho-phonological regularities are simply due to the fact that a finite (albeit large) number of French morphemes and phonemes necessarily fall within one of the two gender categories of the masculine or feminine. However, the multiple spellings associated with morpho-phonological endings along with their highly variable predictability and idiosyncracies (i.e., fluctuating gender, grammatical homonyms, homophones, epicenes, compounds) create a seemingly unacquirable system. So, how do French NSs acquire grammatical gender? What cues do they rely on and do they do so consistently?

REVIEW OF SELECTED L1 FRENCH STUDIES

This section reviews studies that investigated how grammatical gender is acquired, assigned and/or processed by French NSs to address the issue of the types of cues (i.e., morphological and/or phonological) on which they may rely.

---

[11] As pointed by an anonymous reviewer and the literature (e.g., Gross, 1990; Mathieu-Colas, 1996; Savary, 2000) the classification of compounds is of course much more complex than this brief overview.

L1 acquisition studies

Tucker et al. (1977) ran five studies to test NSs' ability in assigning gender to real words – common and rare – and pseudo-words with real phonological and/or morphological endings. The stimuli consisted in randomized list of words previously recorded and presented orally. The first study tested 402 students (8 to 13 years old) and showed that they performed well on endings with high predictability values such as *–ais* [e], *–oi* [wa], *–illon* [i jõ], *–eur* [œr], *–oir* [war];[12] however, they showed some "confusion and ambiguity" (*ibid*: 25) with the endings *–é, –eure* and *–oire* which have low predictibility values (e.g., *–oire* and *–eure* are both 54% masculine). Tucker et al.'s suggestion that "native speakers might disambiguate the total category of "nouns ending with *–é*" by storing separately information about those that end with *–té*, with *–ré*, etc" is difficult to accept as it lacks a principled basis (i.e., aside from gender, it is unclear what the motivation would be), and unnecessarily complicates the processes of lexical storage and retrieval.

Karmiloff-Smith (1979: 167) emphasized the importance of phonological rules as well, contending that French NSs construct "a very powerful, implicit system of phonological rules" that allows them to assign gender to most of the nouns they encounter. Her data showed that at about 5 years old, children prefer indefinite articles as gender cues over phonological noun cues, when these two types of cues were in conflict, while children as young as 3 years old used indefinite articles when they could not rely on phonological cues. The stimuli were presented aurally and the children responded orally.

Seigneuric, Zagar, Meunier and Spinelli (2007) address two limitations of the study conducted by Karmiloff-Smith (1979): first, the fact that only four pseudo-words were used, and second, the possibility that because the children's ability to assign gender was tested with an agreement task (producing a noun phrase with a determiner and an adjective), two different types of ability – assignment and agreement – were confounded. In Seigneuric et al., a large number of children (n=144) from 3- to 9-years old repeated pseudo-words (masculine, feminine, neutral endings) after the researcher, before indicating their gender either by providing a determiner verbally or by pointing to a picture. The children's ability to appropriately assign gender increased with age, with children as young as 4 years old showing they had acquired "a probabilistic system of phonological gender cues that is intimately connected to semantic cues" (*ibid*: 241), replicating findings in French and other languages (Sera, Elie, Forbes, Burch, Rodriguez and Dubois, 2002). However, Seigneuric et al. point out "that showing an influence of word endings in gender attribution to words or pseudo-words does not imply that this type of

---

[12] It is unclear why Tucker et al. (1977) classify [e] as being reliably masculine since, according to their own data, only 57% of tokens ending in [e] are masculine; however, 85% of nouns ending in [wa] are masculine, a much better predictability value.

123

information plays a major role in gender retrieval during "normal" comprehension" (*ibid*: 244).[13]

In Harley ([1979](#)), children in immersion programs experienced greater difficulties in assigning gender during an oral task than monolingual francophones, and overgeneralized the masculine. The same finding was reported in Marinova-Todd ([1994](#)) with older learners, and in Tarone, Frauenfelder and Selinker ([1976](#)) with grade 2 French immersion students. In a longitudinal study, Spilka ([1976](#)) found that immersion students (grades 1 through 6) experienced persistent difficulties with gender, while Taylor-Browne ([1984](#)) reported that other immersion learners did not appear to use phonological, semantic or syntactic cues to assign gender to novel nouns in a variety of experimental tasks.

More recent studies by Boloh and his colleagues (Boloh and Ibernon, [2010](#), [2013](#); Boloh, Escudier, Royer and Ibernon, [2012](#)) found no evidence that French children or adults use a phonological strategy when selecting a gender-marked determiner for inanimate pseudo-nouns in an elicited production task. When 4- to 10-year-old children performed the same task with incongruent determiner-noun pairs of inanimate pseudo-nouns, they never relied on phonological cues, leading Boloh et al. ([2012](#)) to conclude that children use the masculine as a default gender, and that the feminine is acquired based on its co-occurrence with feminine determiners. Boloh and Ibernon ([2013](#)) administered an oral description task with pseudo-words to children (n=182) of various age groups (from 3;8 to 12;6), adolescents (n=24) and young adults (n=22), and similarly report no evidence of a phonological strategy.

Gender processing studies

Gender processing studies suggest that NSs assign gender more rapidly and accurately to words and pseudo-words exhibiting regular gender-ending correspondence than to words with unpredictable endings (e.g., Desrochers and Paivio, [1990](#); Desrochers, Paivio and Desrochers, [1989](#); Holmes and Dejean de la Bâtie, [1999](#); Holmes and Segui, [2004](#), [2006](#); Taft and Meunier, [1998](#)). Gender decisions are also faster when nouns refer to human entities rather than to inanimate objects (Desrochers and Brabant, [1995](#)). However, regular endings do not facilitate gender attribution to vowel-initial nouns, prompting the claim that endings play a secondary role to determiners (Desrochers and Brabant, [1995](#); Desrochers and Paivio, [1990](#); Desrochers et al., [1989](#); Taft and Meunier, [1998](#)). Thus, Holmes and Dejean de la Bâtie ([1999](#): 480) contend that "[g]iven their extremely high frequency of occurrence in the language, definite and indefinite articles would be expected to play a major part in this process", referring to gender acquisition as "purely associative mechanisms". A claim based on empirical evidence showing that NSs "take substantially longer to indicate the gender of isolated nouns when asked

---

[13] Boloh and Ibernon ([2013](#)) also criticize Karmiloff-Smith's study and report that they were unable to replicate her findings.

to use these labels rather than to supply a determiner of the appropriate gender (Desrochers and Paivio, 1990; Desrochers, Paivio and Desrochers, 1989)(*ibid*: 481).

However, written corpus studies suggest that determiners may not be reliable gender markers. The analysis of 5016 DPs (*i.e.*, noun phrases with determiners and/or adjectives as tokens, not types) from a written corpus of newspapers and magazines revealed that only 41.22% of nouns had a gender-marked determiner, while 49.76% of noun tokens were not gender-marked at all (the remaining 9.01% of nouns lacked a gender-marked determiner, but were modified by a gender-marked adjective)(Ayoun 2010). Similarly, Hug (1989) reports that only 42.48% of nouns were gender-marked in a written corpus of 48 literary excerpts. The analysis of two oral corpora – 37.7 hours of immersion teachers in their classrooms and 33 hours of parent–child interactions – allowed Poirier and Lyster (2014) to report that 71% of the direct object clitics *le, la, l', les* used by the teachers, and 54% of the pronouns used by the parents were not gender-marked. In Poirier (2012), the same oral corpus of French Canadian immersion teachers revealed that the gender of 50% of nouns is indicated on their determiners or adjectives, while 50% of nouns are not gender-marked at all.

The written and oral input is thus ambiguous at best, but when it is unambiguous, it is used by NSs according to an experimental study monitoring the eye movements of participants as they performed a picture identification task in Dahan, Swingley, Tanenhaus and Magnuson (2000). Auditory stimuli were played and participants clicked on the relevant picture. Findings show that a gender-marked determiner (as opposed to a non gender-marked determiner) eliminated the activation of gender-inconsistent alternatives, constraining the possibilities entertained for noun recognition.

Likewise, Colé, Pynte and Andriamamonjy (2003) recorded lexical decision times and eye movements while French NSs performed a visual word recognition task with written stimuli targeting two types of regularities: ending-to-gender (i.e., final letters are predictive of gender) and gender-to-ending (e.g.*,* the feminine generally predicts the final letter *e*). The predictability of the nouns' endings was found to facilitate their recognition, at least for low frequency nouns. The authors conclude that NSs use noun ending-gender regularities in word recognition.

Spinelli and Alario (2002) carried out two cross-modal semantic priming experiments with auditory stimuli to determine how the two meanings of homophone words are activated by French NSs. The first experiment that presented words in isolation (a pair of homophones with both related and unrelated target words) showed that both genders are activated, but that there was a priming effect for the primary meaning. In the second experiment, homophones were presented with a determiner marked for the secondary meaning (e.g., *la selle* 'saddle' instead of *le sel* 'salt'),[14] and findings only showed evidence of activation for the meaning triggered by the gender of the determiner. Spinelli and Alario conclude "that

---

[14]Primary and secondary meanings were based on the frequency value of the BRULEX database (Content, Mousty and Radeau, 1990).

a gender-marked context constrains the meanings that are accessed during the perception of homophones. This pattern seems better explained by locating the gender context effect at the level of word form activation. It further suggests that homophone words do not share a common representation at this level" (*ibid*: 467).

Spalek, Franck, Schriefers and Frauenfelder (2008) investigated the role of phonological regularities (i.e., gender is either congruent or incongruent with phonological endings) on gender processing: the first experiment focused on auditory language comprehension, while the second targeted language comprehension, both with auditory stimuli. In the gender decision task of the first experiment, participants were faster with nouns exhibiting congruent endings than incongruent endings, while the effect was weaker in the lexical decision task. No congruency effect was found with the picture naming task administered in the second experiment.[15]

Meunier, Seigneuric and Spinelli (2008) suggest that the noun predictability effect found in these studies may be due to NSs' ability to activate and process the morphological composition of complex nouns. Two experiments were carried out to investigate whether the gender of morphologically complex words derived from a base with an opposite gender is retrieved more slowly than the gender of a noun whose base matches that of the derived noun. The results of a gender assignment task, which presented written stimuli on a computer screen showed that participants decomposed the nouns into their constituent morphemes and activated their morpheme genders to identify them. However, as acknowledged by the authors, it is unclear whether the gender effect was in fact morphemic as opposed to orthographic, and few stimuli (both words and endings) were used.

Muller-Gass, Gonthier, Desrochers and Campbell (2000) monitored the cognitive processing occurring before participants (n=10) indicated their response to stimuli in gender assignment tasks by measuring Event Related Potentials (ERP), in particular the P3 – a late positive wave whose amplitude is reduced with more difficult decisions. Participants viewed the written stimuli on a computer screen and pressed a key to indicate whether a gender label was correct or not while their EEG activity was recorded. Results showed a slower reaction time for superordinate labels (masculine, feminine) than for determiners (*un, une*), as well as for low as opposed to high frequency nouns; however, no significant P3 latency differences were found between determiners and superordinate labels. "Results are consistent with the hypothesis that participants decide on the gender of the target word by constructing a virtual noun phrase and subsequently verify the match between virtual article and gender labels presented on the screen" (*ibid*: 3530).

To sum up this brief literature review and to address the questions asked at the end of Section 2: French NSs appear to acquire grammatical gender by using the most reliable cues available to them, that is, the phonological cues of highly predictable endings, but not necessarily for vowel-initial nouns. More

---

[15] The study is limited by the small sample of nouns and the fact that they were all common words (aside from the experimental pseudo-words).

specifically, high frequency nouns were correlated with faster reaction times and fewer errors than low frequency nouns particularly when their endings were also highly predictive of their gender. Conversely, participants performed poorly with phonological endings that have low predictability. However, it was also found that they use semantic cues and (in)definite articles as well, in addition to being faster with nouns exhibiting congruent endings rather than incongruent endings, as could be expected.

None of these studies systematically investigated common versus uncommon nouns, vowel-initial versus consonant-initial nouns or compounds, nouns exhibiting both masculine and feminine genders, all at once; nor did they ask participants to think about the strategies they may use to assign a gender to stimuli. This is what the present study does to address the questions of how French NSs acquire grammatical gender, what cues do they rely on and whether they do so consistently.

METHODOLOGY

Research questions

The main research questions are as follows: a) how accurate and consistent are French NSs in assigning gender to nouns? b) what strategies do they use when they are unsure of the gender of a noun?

The following predictions will be tested: a) participants will be more accurate and consistent on common than uncommon nouns, animate than inanimate nouns, consonant–initial than vowel initial nouns; b) compounds will be overwhelming masculine; c) there will be a lexical effect in that the participants' performance will vary with the stimuli.

These predictions imply that there should be a lexical effect that derives from the hypothesis that gender must be acquired for each individual lexical item as morpho-phonological cues alone are unreliable and vary greatly; lexical knowledge also comes into play in the case of idiosyncracies (*i.e.*, fluctuating gender, grammatical homonyms, homophones, *épicènes*, compounds) and uncommon nouns.[16]

Participants and tasks

French NSs (n=168) who lived in various cities in France (e.g., Paris, Toulouse, Nantes, Montpellier, Caen) at the time of the data collection were recruited through academic listservs and asked to enlist their friends and families in order to reach people of various socio-economic backgrounds. The resulting composition of the participant pool was graduate students (n=57), professors (n=49), professionals with graduate degrees (n=13), non professionals (n=35), retired (n=14). Non professional participants had graduated from high school, but did not attend college,

---

[16] Following Hawkins and Towell (2010), the stimuli only included compounds written with (or rarely without) a dash.

while professionals had graduate degrees, but were not academics. Most, but not all, of the retired participants indicated that they were former teachers or professors. There were 38 male and 130 female participants who averaged 39.51 years in age (range 19–74). They filled out a written background questionnaire before performing a written gender assignment task: they were asked to decide whether each isolated word was masculine, feminine, both, or if they didn't know. These four options were presented in a pull-down menu. They were also asked to check a box to indicate when they didn't know the meaning of a word.[17] The participants also completed a preference/grammaticality judgment task and an acceptability judgment task as part of a larger study; these findings will be reported elsewhere. All the tasks and stimuli were written and accessible from a webpage. The study was composed of three sessions to accommodate the data collection of two different tasks per session and to avoid performance errors due to fatigue, boredom or lack of concentration, as well as to include a large number of stimuli for each of the two different tasks administered per session. The stimuli were thus randomly divided across three sessions, but there were no differences in the way the gender assignment task was administered. In other words, participants performed three times the same task, once every two days, with different stimuli.

The stimuli were composed of a total 295 individual words illustrating vowel- and consonant-initial nouns, compounds, common and uncommon nouns, the former being defined as high frequency nouns (based on lexical databases such as Content et al., 1990 and lexique.org) such as *mort* 'death' or *chose* 'thing', as opposed to the latter which are much less frequent in daily use such as *uvule* 'uvula' or *renoncule*; 'buttercup' as well as nouns whose gender is known to be difficult even for NSs, but are common (according to Holmes and Dejean de la Bâtie, 1999; Batchelor and Offord, 2000; Ayoun, 2007).[18] Grammatical homonyms and épicènes were included so that the stimuli would be representative of French vocabulary. The distribution of the stimuli is displayed in Table 3 (see also Appendix A).

At the end of the last session, the participants were asked to think about the strategies they may have used when they were unsure of the gender of the nouns.

---

[17]No reaction time was recorded as it was not deemed relevant to the task and may even affect the findings if the participants feel pressured or pressed for time. Moreover, different reaction times can be attributed to individual differences in working memory, not in competence, and present various methodological challenges such as accurate measurements. The participants were instructed as follows: *Veuillez indiquer si les noms suivants sont masculin, féminin, ou les deux à la fois. Si vous ne savez pas, choisissez cette option. Si vous ne connaissez pas le(s) sens du mot, veuillez l'indiquer aussi en cochant la case à côté de 'sens inconnu' et essayez quand même de choisir une option pour le genre du nom. Merci!* 'Please indicate whether the following nouns are masculine, feminine, or both. If you don't know, choose that option. If you do not know the meaning(s) of the word, please indicate that as well by clicking in the box next to 'unknown meaning' and try to choose an option for the gender of the noun anyway. Thank you!' [translation mine]

[18]The stimuli include 54 nouns (30 masculine and 24 feminine) among the common nouns whose gender is known to be difficult from Batchelor and Offord (2000) such as *abîme* 'abyss', *chaume* 'thatch', *effluve* 'smell'.

Table 3. *Distribution of the stimuli*

	session 1	session 2	session 3	total
masculine	20	42	62	**124**
feminine	15	14	28	**57**
both	60	41	13	**114**
total	**95**	**97**	**103**	**295**
simple	76	68	77	221
compound	19	29	26	74
common	83	88	80	251
uncommon	12	9	23	44
vowel initial	21	15	30	66
consonant initial	74	82	73	229
animate	41	34	22	97
inanimate	54	38	81	173
(in)animate	5	15	0	20

Table 4. *Study design*

	Task 1	Task 2
Session 1	gender assignment task (95 stimuli)	acceptability judgment task
Session 2	gender assignment task (99 stimuli)	preference/grammaticality judgment task
Session 3	gender assignment task (102 stimuli) + strategies used	preference/grammaticality judgment task

They were given a choice of six different strategies, and for each one indicated whether they used it 'always', 'often', 'sometimes', 'rarely' or 'never'. The strategies were chosen to test hypotheses proposed from the literature reviewed above, that is whether native speakers use determiners and/or phono-morphological cues, and/or whether they go beyond the noun itself and exploit any indications given by the adjective or the sentential context. The latter was based on Gollan and Frost's (2002: 642) claim that "gender is accessed more efficiently in the presence of syntactic context (Roelofs, 1992; Schriefers, 1993)". No time limit was placed on performing the tasks. When participants were finished, they clicked on the submit button. The raw data were saved to a folder online to be later coded and analyzed with SPSS. The study design is summarized in Table 4.

FINDINGS

Overall accuracy

The accuracy means displayed in Table 5 show how often participants assigned the appropriate gender to the stimuli.

129

Table 5. *Accuracy on gender assignment tasks*

		Participants' responses		
		correct	incorrect	don't know
masculine	count	17160	3672	689
	% within gender	**82.4%**	**17.6%**	**3.3.%**
feminine	count	6946	2462	323
	% within gender	**73.8%**	**26.2%**	**3.4%**
both masculine and feminine	count	12086	7570	323
	% within gender	**61.5%**	**38.5%**	**1.6%**
	total	36192	13704	1335
		72.5%	**27.5%**	**2.7%**

Table 6. *(Un)known meaning/gender accuracy*

			Incorrect	Correct
meaning/gender	known meaning, gender attributed	count	12296	35986
		% within meaning	25.5%	**74.5%**
	unknown meaning, gender attributed	count	146	175
		% within meaning	45.5%	**54.5%**
	unknown meaning, unknown gender	count	505	12
		% within meaning	97.7%	**2.3%**
	known meaning, unknown gender	count	757	19
		% within meaning	97.6%	**2.4%**
	Total	count	13704	36192
		% within meaning	27.5%	**72.5%**

Participants obtained an overall accuracy of 72.5% and their performance was significantly different with the gender of the stimuli (Pearson $\chi^2 = 2834.692$, df $= 4$, $p<.001$): 82.4% for masculine nouns, 73.8% for feminine nouns and 61.5% for nouns which are both masculine and feminine. They chose the option 'I don't know' for a very small percentage of the stimuli (3.3% masculine, 3.4% feminine, 1.6% both masculine and feminine), suggesting they were confident. They could also indicate whether they knew the meaning of the stimuli or not. These findings are displayed in Table 6.

Participants indicated not knowing either the meaning or the gender for only 2.3% of the stimuli; they knew the meaning, but didn't know the gender of only 2.4% of the stimuli. When they didn't know the meaning of the noun, they still correctly attributed a gender for 54.5% of the stimuli. And finally, when they knew the meaning of the stimuli, they were accurate in attributing the gender of these stimuli most of the time (74.5%). These findings are statistically significant (Pearson $\chi^2 = 3341.847$, df $= 3$, $p<.001$). So a strong confidence indicated by

Table 7. *Accuracy on GAT by variables*

		accuracy	Pearson χ^2	df	*p*
commonality	common	73.6%	146.106	1	.001
	uncommon	67.2%			
animacy	animate	78.2%	986.696	2	.001
	inanimate	65.2%			
	both (in)animate	67.8%			
initial	vowel	67.6%	177.783	1	.001
	consonant	74.0%			
suffix	yes	77.6%	131.521	1	.001
	no	71.5%			
compound	no	68.4%	1412.152	2	.001
	yes, animate	73.6%			
	yes, inanimate	86.7%			

low percentages for unknown meaning and gender (2.3% and 2.4%, respectively) did not lead to a solid performance (74.5%), well below the 90% criterion usually required of native speakers (e.g., Dronjic and Helms-Park, 2014).

Accuracy by variables

The data were also analyzed based on several variables such as commonality, suffix and initials; the accuracy percentages and statistics are displayed in Table 7.

Participants performed better on common than uncommon words (average of 73.6% vs. 67.2%), animate than inanimate nouns or both animate and inaninate nouns (average of 78.2%, 65.2%, 67.8%, respectively), as well as consonant-initial nouns versus vowel-initial nouns (74.0% vs. 67.6%). They were slightly more accurate with the stimuli, which had a suffix (77.6% vs. 71.5%), and they obtained better results with compounds than non-compounds (68.4%) whether they were animate (73.6%) or inanimate (86.7%). All of these findings are statistically significant. However, no correlation was found between the frequency of the stimuli as determined by lexique.org and the participants' performance (Pearson correlation = –0.067, *p* = 0.246).[19]

More detailed findings regarding the compounds are displayed in Tables 8a and 8b.

Participants performed significantly better on compounds (91.2%) than non-compounds (from 63.0% to 73.7%). They also did better with masculine stimuli (both compounds and non-compounds) than feminine stimuli, or stimuli that

---

[19]Lexique.org is an online corpus that contains 14.7 million words from literary texts published between 1950 and 2000 (New, Pallier, Ferrand and Matos, 2001). The latest version includes compounds and the number of homophones, but with a single frequency index which diminishes its reliability.

Table 8a. *Accuracy on compounds*

| | | participants' responses | | | |
	gender	M	F	MF	don't know
non compounds	M	**73.7%**	12.5%	9.3%	4.4%
	F	10.9%	**72.9%**	12.8%	3.4%
	MF	17.1%	18.4%	**63.0%**	1.5%
inanimate compounds	M	**91.3%**	3.8%	2.6%	2.3%
	F	10.3%	**85.6%**	0.7%	3.4%
	MF	58.6%	11.8%	**23.5%**	6.1%
animate compounds	M	**91.2%**	0.5%	7.5%	0.8%
	MF	38.2%	0.8%	**59.0%**	2.0%

Table 8b. χ^2 *tests for compounds*

χ^2 tests		Value	df	Asymp. Significance (2-sided)
non compounds	Pearson χ^2	22931.728	6	.001
yes, inanimate	Pearson χ^2	5803.139	6	.001
yes, animate	Pearson χ^2	553.136	3	.001

Table 9. *Accuracy on épicènes and grammatical homonyms*

| | | participants' responses | | | |
	gender	M	F	MF	don't know
epicenes	MF	18.0%	1.7%	**78.8%**	1.4%
animate-inanimate homonyms	MF	21.6%	25.5%	**49.5%**	3.3%
inanimate-inanimate homonyms	MF	19.4%	35.3%	**44.0%**	1.3%

were both masculine and feminine. This is particularly noticeable for inanimate compounds that were found to be masculine (58.6%) more often than both masculine and feminine (23.5%), and to a lesser extent for animate compounds (38.2% masculine vs. 59.0% masculine and feminine). The percentages for 'don't know' vary and are the highest for inanimate compounds (6.1%), but remain low overall.

Table 9 displays the findings for *épicènes* and grammatical homonyms, which are all both masculine and feminine.

The participants performed much better on epicenes whose referents are animate (e.g., *adversaire* 'adversary'-FEM-MSC) than on homonyms with both an animate referent and an inanimate referent (e.g., *aide* 'aid, assistant'-FEM-MSC/'help'-FEM) or homonyms with two inanimate referents (e.g., *vase* 'vase'-MSC/'mud'-FEM). Their accuracy percentages are below chance level for the homonyms (49.5%
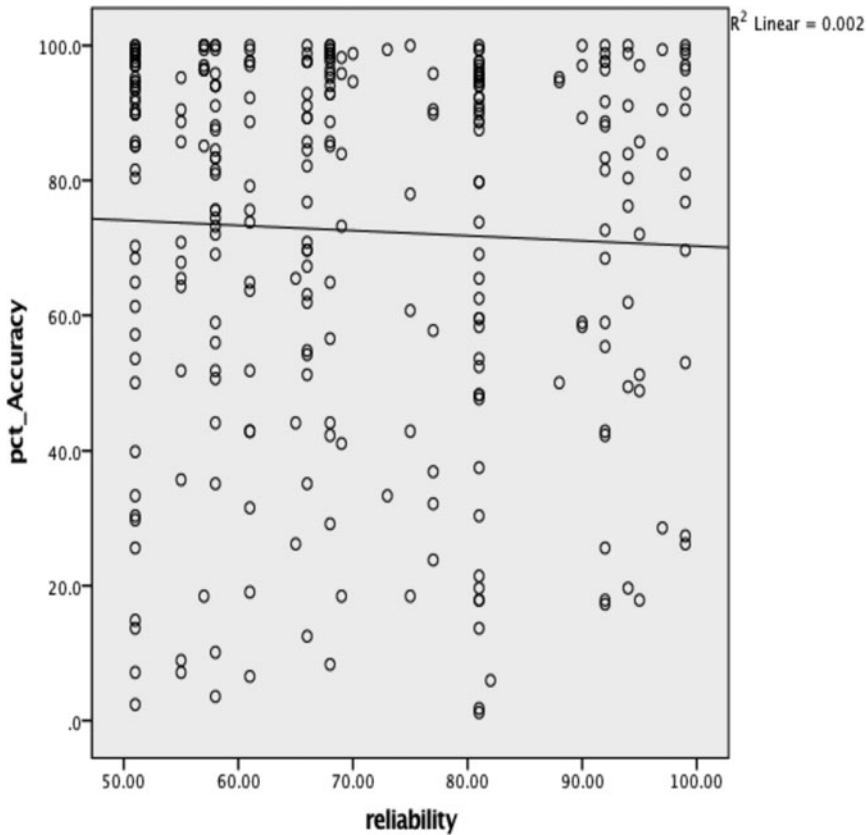
Figure 1. *Correlation accuracy and reliability index*

and 44.0%) and higher for the epicenes (78.8%), but statistically significant (Pearson $\chi^2 = 3590.864$, df $= 6$, $p = 0.001$), although the 'don't know' percentages are very low, indicating a strong confidence level.

Another interesting finding is the lack of correlation between the participants' accuracy in attributing gender to the stimuli and the reliability index of the phonological endings (Tables 1 and 2). As Figure 1 shows, there was no correlation (Pearson correlation $= -0.042$, $p = 0.474$).

Finally, although the findings for accuracy by profession are statistically significant as shown in Table 10, the overall accuracy means are not very far apart from the highest means for retired participants who were mostly professors (76.4%) to the lowest for graduate students (70.4%).

Given the high level of education of the participants, findings may have been worse with participants with lower levels of education, that is, below the *baccalauréat* (high school diploma). There is no statistical difference between the performance of

133

Table 10. *Accuracy by profession and sex*

		n	overall accuracy	Pearson χ^2	df	*p*
Profession	professor	49	74.0%			
	graduate students	57	70.4%			
	retired	14	76.4%	88.293	4	.001
	professionals with graduate degree	13	73.3%			
	non professionals	35	72.2%			
sex	female	130	72.4%	1.658	1	.200
	male	38	73.0%			

male vs female participants which is not surprising given the disparity between the number of male vs female participants, but their sex is unlikely to be a significant factor in their performance anyway.

Findings by stimuli

Table 11 displays stimuli that got low accuracy means. There are first a few compounds and then non-compounds by increasing means. Some of the compounds are common and yet were rarely attributed the appropriate gender (e.g., *mi-temps* 'part-time, half-time' 26.2%, *après-guerre* 'post-war period' 37.5%, *garde-à-vue* 'detention' 51.2%), while others are more rare and clearly affected the participants' performance (e.g., *garde-cendre* 'ash guard' 1.2%, *porte-voix* 'megaphone' 6.0%, *micro-cravate* 'lapel microphone' 57.1%).

Among the non-compounds, there are common (e.g., *amour* 'love' 59.5%, but *amours* 'loves' 47.6%, *armistice* 'armistice' 51.8%, *primeurs* 'early vegetables' 19.6%), uncommon nouns (e.g., *renoncule* 'buttercup' 44.0%, *adobe* 'adobe' 26.2%, *effluve* 'scent' 18.5%) as well as quite a few grammatical homonyms (e.g., *espace* 'space' 19.0%, *vase* 'vase, mud' 58.3%, *somme* 'nap, addition' 58.9%). The accuracy means go from 1.2% to 59.5%, indicating a strong lexical effect.

Strategies

At the end of the last session, participants were asked to indicate whether and how often they may have used several possible strategies when they were unsure of the gender of a noun.

The percentages displayed in Table 12 show that the NSs did not favor a single strategy since they often selected more than one. The first four were chosen the most often and are presented in decreasing order: using the word with an adjective (54.8% combining 'always' and 'often'), using the word with an article (45.9%), using the word in a sentence and sounding it out mentally (42.3% for both strategies). What they did not do was use the first few letters (91.7% combining 'rarely' and 'never')

Table 11. *Findings by stimuli*

Stimuli	gender	accuracy	stimuli	gender	accuracy
garde-cendre	M	1.2%	crêpe	MF	29.8%
porte-voix	M	6.0%	termite	M	30.4%
mi-temps	MF	26.2%	enseigne	MF	31.5%
après-guerre	MF	37.5%	greffe	MF	32.1%
avant-guerre	MF	48.2%	vague	MF	33.3%
garde-à-vue	F	51.2%	uvule	F	35.1%
micro-cravate	M	57.1%	astérisque	M	35.7%
ayant cause	M	58.9%	cartouche	MF	35.1%
ombre	MF	1.8%	cave	MF	41.1%
bugle	M	3.6%	boum	MF	42.3%
délices	F	6.5%	mime	MF	42.9%
carpe	MF	7.1%	casse	MF	42.9%
Pâques	MF	7.1%	auspices	M	42.9%
basque	MF	8.9%	faune	MF	44.0%
hymne	M	8.3%	amours	F	47.6%
finale	MF	10.1%	superbe	F	44.0%
barde	MF	12.5%	renoncule	F	44.0%
gite	MF	13.7%	membre	MF	48.2%
œuvre	F	13.7%	orbite	F	50.0%
trompette	MF	14.9%	argile	F	51.8%
crème	MF	17.3%	armistice	M	51.8%
pantomime	MF	17.9%	plastique	MF	51.8%
chèvre	MF	17.9%	aérogare	F	52.4%
merci	MF	18.5%	chaume	M	55.4%
effluve	M	18.5%	parallèle	MF	56.0%
apogée	M	18.5%	épitaphe	F	57.7%
espace	MF	19.0%	pupille	MF	56.5%
primeurs	F	19.6%	en-tête	M	53.6%
primeur	F	21.4%	antipode	M	54.2%
coche	MF	23.8%	solde	MF	54.8%
geste	MF	25.6%	manœuvre	MF	58.3%
adobe	M	26.2%	idole	F	58.9%
gens	MF	27.4%	somme	MF	58.9%
radio	MF	28.6%	vase	MF	58.3%
			amour	M	59.5%

or the last few letters (78.8%), contra Tucker et al. (1977). So phonology played a role (sounding out the words was used 42.3% of the time), but the gender provided by the determiner such as *un*/*une* 'a'-MSC/FEM or *le*/*la* 'the'-MSC/FEM (presumably as an associative process) did as well (as found in Muller-Gass et al., 2000). This finding supports earlier claims that "when asked to classify a word's gender, people try to evoke implicitly the closest lexical associate, typically the definite article" (Holmes and Segui, 2004: 428). The participants who chose 'did something else' rarely provided an explanation. The few explanations participants wrote are about creating a context around the word, using general world knowledge, analogy or etymology (see Appendix B). The explanations provided appear to be typical of declarative

Table 12. *Strategies used to attribute gender*

	always	often	sometimes	rarely	never
Used the word with an article	28.6%	17.3%	25.0%	29.2%	0.0%
	45.9%			29.2%	
Used the word with an adjective	12.5%	42.3%	26.8%	10.1%	8.3%
	54.8%			18.4%	
Tried to use the word in a sentence	24.4%	17.9%	21.4%	11.9%	24.4%
	42.3%			36.3%	
Said it in my head to sound it out	28.6%	13.7%	23.8%	30.4%	3.6%
	42.3%			34.0%	
Used the first few letters	3.6%	2.4%	2.4%	1.2%	90.5%
	6.0%			**91.7%**	
Used the last few letters	10.1%	7.1%	4.8%	10.7%	67.3%
	17.2%			**78.0%**	
Did something else	9.5%	6.5%	2.4%	3.6%	78.0%
	16.0%			**81.6%**	

knowledge as participants were able to articulate metalinguistic knowledge such as 'I used similar compound words, or tried to see if the first word in the compound words were verbs or nouns'; in this case, the participant presumably thought that the type of compound influenced its gender. Two other explanations – 'I used a past participle with a phonetically realized feminine agreement' and 'When I really didn't know, I tried to decompose the word and to guess its etymology'– also indicate a certain amount of metalinguistic knowledge, which is not surprising given that almost a third of the participants were teachers/professors (n = 49). If the participants tried to recall similar lexical items, they were using their procedural memory system as well as argued by Ullman (2004: 233) according to whom "lexical memory depends largely on the declarative memory system whereas aspects of grammar depend on the procedural memory system". To remember idiosyncratic lexical information, they probably relied on their mental lexicon (e.g., Chomsky, 1995). Unfortunately, the participants did not provide enough explanations to better characterize their strategies and we can only speculate as to the exact type of knowledge they used, consciously or unconsciously, to make a decision while performing this task, especially given empirical findings that native speakers are not immune to task effects any more than L2 learners are (e.g., Foster and Tavakoli 2009).

DISCUSSION AND CONCLUSION

The present study set out to determine how reliable and consistent French NSs are at assigning the appropriate gender to common, uncommon nouns and compounds in a written gender assignment task and what strategies they may use in doing so. It was hypothesized that NS intuition and purely associative mechanisms are not sufficient; instead, gender must be acquired for each individual lexical item as morpho-phonological cues alone are unreliable, and numerous idiosyncrasies create ambiguities.

The results of a gender assignment task revealed strong lexical and gender effects with an overall accuracy of 72.5% and a significantly better performance on masculine nouns (82.4%) than feminine nouns (73.8%) or nouns which are both masculine and feminine (61.5%). The NS speakers' performance also varied depending on whether the stimuli were compounds or not, common or uncommon, or had a vocalic or consonantal initial. A strong lexical effect was also found, confirming our hypothesis that gender must be acquired for each individual lexical item. Following Acquaviva (2008, 2009) for Italian pairs such as *braccio ∼ braccia* 'arm ∼ arms' with different gender values in singular and plural, French grammatical homonyms may be considered as distinct lexemes as argued by Caramazza, Costa, Miozzo and Bi (2001) based on a translation task with Spanish–English bilingual speakers who needed more time to translate the low frequency twin in a pair of homophones. Biedermann, Blanken and Nickels (2002) also argue that although homophones may share a single word form representation for the sake of an efficient speech-processing system, when they have different grammatical genders, they should have different lemmas. Spinelli and Alario (2002) reach the same conclusion based on reaction times in the production of French homophones, supporting the Independent Network Model (Caramazza and Miozzo, 1997).[20]

The answer to our second research question related to the strategies French NSs reported using which varied from using the word with an adjective (54.8%), an article (45.9%), and in a sentence or sounding it out mentally (42.3% for both strategies). However, they did not use the last few letters (78.8%), contra Tucker *et al.* (1977).

So phonology played a role, but so did the gender provided by the determiner. It thus appears that associative processes are insufficient (contra Sagarra and Herschensohn, 2010) when NSs are confronted with uncommon nouns, particularly if they are feminine and vowel-initial. Morpho-phonological cues alone are clearly unreliable – practically useless in the case of compounds – and lexical knowledge is essential in the case of idiosyncracies (e.g., Ayoun 2007).

Holmes and Dejean de la Bâtie (1999: 490) found that French NSs "took significantly longer to classify pseudo-words than regular words" and "unexpectedly [. . .] also made substantially more errors classifying pseudo-words than regular words", but appeared to have used sublexical associations in attributing gender. Moreover, they reported that NSs "did not classify the items in terms of their endings 20% of the time (*ibid*: 493). A second experiment with pseudo-words did not support Tucker et al.'s (1977) claim that NSs systematically establish a correspondence between endings and gender.

A task effect may explain some of these low accuracy percentages: participants were asked to attribute a gender to individual, decontextualized words from a pull-down menu (M, F, MF, DK). Presenting the words in context would clearly illustrate

---

[20] Another view holds that homophones share a phonological representation, but have different semantic and grammatical representations (e.g., Jescheniak and Levelt, 1994; Levelt, Roelofs and Meyer, 1999).

the nouns' meaning and thus activate both genders in the case of grammatical homonyms if two different stimuli were used to illustrate the two different meanings (e.g., *la vase* 'mud' vs *le vase* 'vase'). It would also be interesting to investigate whether the same gender assignment task carried out with auditory stimuli would yield similar results than with written stimuli. The gender processing studies reviewed above suggest that there may be differences since "the stimulus information in a written noun is processed in parallel, auditory information is processed sequentially" (Spalek et al., 2008: 431).

Grammatical gender is a complex lexical property that poses retrieval difficulties even for NSs as evidenced by their inconsistent performance on a gender assignment task with isolated stimuli, thus dispelling the myth of an infallible native speaker. Being at the interface of morpho-phonology and the lexicon, grammatical gender is inherently difficult, but it also appears to be one of the most deeply linguistically entrenched categories (Corbett, 1991; Comrie, 1999), relevant to both the lexicon and morpho-syntax for agreement. It is also important for "disambiguation and reference tracking in discourse (Corbett, 1991).[21]

Without going as far as claiming that the native speaker is dead (Braine, 1999; Edge, 2006; Holliday, 2008), it is important to point out that the idealization of the native speaker cannot be maintained. Native speakers are not immune to performance difficulties due to linguistic complexities, nor are they isomorphic with the standard form of a language, contrary to the academic construct of the native speaker (Davies, 2013). Moreover, their socio-economic and sociolinguistic background does affect their performance (Dronjic and Helms-Park, 2014; Mulder and Hulstijn, 2011).

This conclusion has implications for L2 learners of French as well: acquiring grammatical gender will be a long and arduous process that will require a strong morpho-phonological and lexical knowledge as evidenced by mixed findings. The performance of instructed L2 learners is often inconsistent (e.g., Ayoun, 2007; Dewaele and Véronique, 2001), although form-focused instruction appears to be more efficient than corrective feedback with immersion students (Lyster, 2004), but all participants improved in correctly assigning grammatical gender to French nouns regardless of the type of feedback they received (prompts or recasts) in Lyster and Izquierdo (2009). It is even claimed that L2 learners cannot fully acquire gender features lacking in their L1 (e.g., Franceschina, 2005; Sabourin, 2001; Sabourin, Stowe and Haan, 2006), but L2 learners' performance improves with their proficiency levels (e.g., Sagarra and Herschensohn, 2012), suggesting they may eventually fully acquire grammatical gender although it is low in communicative value (Warden, 1997). The fact that even bilingual children in immersion settings initially have difficulties, but show an improvement in proficiency (e.g., Swain, 2000) indicates that although it is possible for L2 learners to acquire grammatical gender, it will take longer than in L1 acquisition. Finally, in assessing L2 learners we

---

[21] As suggested by an anonymous reviewer, the functionality of grammatical gender may be questioned and it may eventually disappear as argued in Trudgill (1999).

cannot require a level of performance that native speakers do not achieve, especially if we insist on using an idealized native speaker as a yardstick which may doom L2 learners to failure regardless of how near native they may become, as argued by others (e.g., Joseph, 2017).

Address for correspondence:
e-mail: ayoun@email.arizona.edu

REFERENCES

Acquaviva, P. (2008). *Lexical Plurals*. Oxford: Oxford University Press.

Acquaviva, P. (2009). The structure of the Italian declension system. In: F. Montermini, G. Boyé and J. Tseng (eds.), *Selected Proceedings of the 6th Décembrettes*, Somerville, MA: Cascadilla Press, pp. 50–62.

Anastassiadis-Syméonidis, A. (2005). L'attribution du genre aux emprunts du grec moderne au français et à l'anglais américain. In: T. Arnavielle (ed.), *Langues: Histoire et usages dans l'aire méditerranéenne*. Paris: L'Harmattan, pp. 111–122.

Anastassiadis–Syméonidis, A. and Nikolaou, G. (2011). L'adaptation morphologique des emprunts néologiques: en quoi est-elle précieuse? *Langages*, 183: 119–132.

Antes, T. (1993). The effect of enhanced input on the acquisition of phonological gender markers in first year French. PhD thesis, Cornell University.

Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. In: D. Ayoun (ed.), *French Applied Linguistics*. Amsterdam: John Benjamins, pp. 130–170.

Ayoun, D. (2009). The (in)stability of gender assignment and agreement among French native speakers. Talk given at AAAL, Denver.

Ayoun, D. (2010). Corpus data. Shedding the light on grammatical gender . . . or not. *EUROSLA Yearbook*, 10: 119–141.

Ayres-Bennett, W. and Carruthers, J. (2001). *Studies in the Modern French Language: Problems and Perspectives*. Harlow: Longman.

Batchelor, R. E. and Offord, M. H. (2000). *Using French*. 3rd ed. Cambridge: Cambridge University Press.

Beckett, M. (2010). *Gender Assignment and Word-final Pronunciation in French: Two semantic systems*. Munich: Lincom.

Bidot, E. (1925). *La clef du genre des substantifs français (méthode dispensant d'avoir recours au dictionnaire)*. Poitiers: Imprimerie Nouvelle.

Biedermann, B., Blanken, G. and Nickels, L. (2002). The representation of homophones: evidence from remediation. *Aphasiology*, 16: 115–136.

Boloh, Y., Escudier, F., Royer, S. and Ibernon, L. (2012). Determiners and nouns' endings in French children's gender attribution choices. *First Language*, 32: 398–410.

Boloh, Y. and Ibernon, L. (2010). Gender attribution and gender agreement in 4- to 10-year-old French children. *Cognitive Development*, 25: 1–25.

Boloh, Y. and Ibernon, L. (2013). Natural gender, phonological cues and the default grammatical gender in French children. *First Language*, 33: 449–468.

Braine, G. (1999). *Non-native Educators in English Language Teaching*. Mahwah, NJ: Erlbaum.

Caramazza, A., Costa, A., Miozzo, M. and Bi, Y. (2001). The specific-word frequency effect: implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27: 1430–1450.

Caramazza, A. and Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, 69: 231–241.

Carroll, S. (1989). Second-language acquisition and the computational paradigm. *Language Learning*, 39: 535–594.

Carroll, S. (2005). Input and SLA: adults' sensitivity to different sorts of cues to French gender. *Language Learning*, 55, Suppl.1: 79–138.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Clark, E. V. (1985). The acquisition of Romance, with special reference to French. In: D. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition: 1. The Data*. Hillsdale, NJ: Erlbaum, pp. 687–782.

Colé, P., Pynte, J. and Andriamamonjy, P. (2003). Effect of grammatical gender on word recognition: evidence from lexical decision and eye movement experiments. *Perception & Psychophysics*, 65: 407–419.

Comrie, B. (1999). Grammatical gender systems: a linguist's assessment. *Journal of Psycholinguistic Research*, 28: 457–466.

Content, A., Mousty, P. and Radeau, M. (1990). Brulex. Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année Psychologique*, 90: 551–566.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Corbett, G. (2003). Agreement: terms and boundaries. In: W. Griffin (ed.), *The Role of Agreement in Natural Language: Proceedings of the 2001 Texas Linguistic Society Conference, Austin, Texas 2–4 March 2001*. Austin, TX: Texas Linguistic Society, pp. 109–122.

Cyr, M. and Shi, R. (2013). Development of abstract grammatical categorization in infants. *Child Development*, 84: 617–629.

Dahan, D., Swingley, D., Tanenhaus, M. and Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, 42: 465–480.

Davies, A. (2013). *Native Speakers and Native Users*. Cambridge: Cambridge University Press.

Desrochers, A. (1986). Genre grammatical et classification nominale. *Revue Canadienne de Psychologie*, 40: 224–250.

Desrochers, A. and Brabant, M. (1995). Interaction entre facteurs phonologiques et sémantiques dans une épreuve de catégorisation lexicale. *Revue Canadienne de Psychologie*, 49: 240–262.

Desrochers, A. and Paivio, A. (1990). Le phonème initial des noms inanimés et son effet sur l'identification du genre grammatical. *Canadian Journal of Psychology*, 44: 44–57.

Desrochers, A., Paivio, A. and Desrochers, S. (1989). L'effet de la fréquence d'usage des noms inanimés et de la valeur prédictive de leur terminaison sur l'identification du genre grammatical. *Canadian Journal of Psychology*, 43: 62–73.

Deutsch, A. and Dank, M. (2009). Conflicting cues and competition between notional and grammatical factors in producing number and gender agreement: evidence from Hebrew. *Journal of Memory and Language*, 60: 112–143.

Dewaele, J.-M. and Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. *Bilingualism*, 4: 275–297.

Dronjic, V. and Helms-Park, R. 2014. Fixed choice word-association tasks as second language lexical tests: what native-speaker performance reveals about their potential weaknesses. *Applied Psycholinguistics*, 35: 193–221.

Edge, J. (2006). *(Re)locating TESOL in an Age of Empire*. London: Palgrave Macmillan.

Enger, H-O. (2009). The role of core and non-core semantic rules in gender assignment. *Lingua*, 119: 1281–1299.

Foster, P. and Tavakoli, P. (2009). Native speakers and task performance: comparing effects on complexity, fluency and lexical diversity. *Language Learning*, 59: 866–896.

Franceschina, F. (2005). *Fossilized Second Language Grammars: The Acquisition of Grammatical gender*. Amsterdam: John Benjamins.

Gabriel, U. (2008). Language policies and in-group favouritism: the malleability of the interpretation of generically intended masculine forms. *Social Psychology*, 39: 103–107.

Gervais, M-M. (1993). Gender and language in French. In: C. Sanders (ed.), *French Today: Language in its Social Context*. Cambridge: Cambridge University Press, pp. 121–138.

Gollan, T. H. and Frost, R. (2002). Two routes to grammatical gender: evidence from Hebrew. *Journal of Psycholinguistic Research*, 30: 627–651.

Gottlob, R. G., Goldinger, S. D., Stone, G. O. and Van Orden, G. C. (1999). Reading homographs: orthographic, phonologic and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, 25: 561–574.

Gross, G. (1990). Définition des noms composés dans un lexique-grammaire. *Langue Française*, 87: 84–90.

Guilford, J. (1994). Les emprunts à l'anglais dans la presse française pop-rock. PhD thesis, Université Paris-Descartes, Paris V.

Guilford, J. (1999). L'attribution du genre aux emprunts à l'anglais. *La Linguistique*, 35: 65–85.

Gygax, P. and Gabriel, U. (2008). Can a group of musicians be composed of women? Generic interpretation of French masculine role names in absence and presence of feminine forms. *Swiss Journal of Psychology*, 67: 141–153.

Gygax, P., Sarrasin, O., Lévy, A., Sato, S. and Gabriel, U. (2013). La représentation mentale du genre pendant la lecture: état actuel de la recherche francophone en psycholinguistique. *Journal of French Language Studies*, 23: 243–257.

Harley, B. (1979). French gender 'rules' in the speech of English-dominant, French-dominant and monolingual French-speaking children. *Working Papers in Bilingualism*, 19: 129–156.

Hawkins, R. and Towell, R. (2010). *French Grammar and Usage*. 3rd ed. London: Routledge.

Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A. and Schmitz, M. (2004). Functional elements in infants' speech processing: the role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5: 341–353.

Holliday, A. (2008). Standards of English and politics of inclusion. *Language Teaching*, 41: 119–130.

Holmes, V. M. and Dejean de la Bâtie, B. (1999). Assignment of grammatical gender by native speakers and foreign learners of French. *Applied Psycholinguistics*, 20: 479–506.

Holmes, V. M. and Segui, J. (2004). Sublexical and lexical influences on gender assignment in French. *Journal of Psycholinguistic Research*, 33: 425–457.

Holmes, V. M. and Segui, J. (2006). Assigning grammatical gender during word production. *Journal of Psycholinguistic Research*, 35: 5–30.

Hug, M. (1989). *Structure du syntagme nominal français. Etude statistique*. Paris: Champion-Slatkine.

Hulk, A. and Tellier, C. (1998). Conflictual agreement in Romance nominals. In: J.-M. Authier, B. Bullock and L. Reed (eds), *Formal Perspectives on Romance Linguistics*. Amsterdam: John Benjamins, pp. 179–195.

Hulk, A. and Tellier, C. (2000). Mismatches: Agreement in qualitative constructions. *Probus*, 12: 33–65.

Irmen, L. and Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology*, 57: 367–375.

Irmen, L. and Schumann, E. (2011). Processing grammatical gender of nouns: further evidence from eye movements. *Journal of Cognitive Psychology*, 23: 998–1014.

Jescheniak, J. D. and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20: 824–843.

Joseph, J. E. (2017). Extended/distributed cognition and the native speaker. *Language & Communication*, 1–11.

Karmiloff-Smith, A. (1979). *A Functional Approach to Child Language*. Cambridge: Cambridge University Press.

Kastovsky, D. (2000). Inflectional classes, morphological restructuring, and the dissolution of Old English grammatical gender. In: B. Unterbeck, T. Nevalainen and M. Saari (eds.), *Gender and Cognition*. Berlin: Mouton de Gruyter, pp. 709–727.

Kilarski, M. (2001). Gender assignment in English loan-words in Danish, Swedish and Norwegian. PhD thesis, Adam Mickiewicz University.

Levelt, W. J. M., Roelofs, A. and Meyer, A. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22: 1–75.

L'Huillier, M. (1999). *Advanced French Grammar*. Cambridge: Cambridge University Press.

Lyster, R. (2006). Predictability in French gender attribution: a corpus analysis. *Journal of French Language Studies*, 16: 69–92.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26: 399–432.

Lyster, R. and Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 52: 453–498.

MacWhinney, B. (2000). Lexicalist connectionism. In: P. Broeder and J. Murre (eds.), *Models of Language Acquisition: Inductive and Deductive Approaches*. Cambridge, MA: MIT Press, pp. 9–32.

Maratsos, M. and Chalkley, M. A. (1980). The internal language of children's syntax. The ontogenesis and representation of syntactic categories. In: K. E. Nelson (ed.), *Children's Language* (Vol. 2). New York: Gardner, pp. 127–214.

Marinova-Todd, S. (1994). The critical period in second language acquisition: the case of gender. Honors thesis, York University.

Mathieu-Colas, M. (1996). Essai de typologie des noms composés français. *Cahiers de Lexicologie*, 71–125.

Melançon, A. and Shi, R. (2015). Representations of abstract grammatical feature agreement in young children. *Journal of Child Language*, 42: 1–15.

Mel'čuk, I. A. (1958). Statistics and the relationship between gender of French nouns and their endings. In: V. J. Rozencvejg (ed.), *Essays on Lexical Semantics* I. Stockholm: Skriptor, pp. 11–42.

Meunier, F., Seigneuric, A. and Spinelli, E. (2008). The morpheme gender effect. *Journal of Memory and Language*, 58: 88–99.

Mulder, K. and Hulstijn, J. (2011). Linguistic skills of adult native speakers, as a function of age and level of education. *Applied Linguistics*, 32: 475–494.

Müller, N. (1990). Developing two gender assignment systems simultaneously. In: J. Meisel (ed.), *Two First Languages: Early Grammatical Development in Bilingual Children.* Dordrecht: Foris, pp. 193–234.

Muller-Gass, A., Gonthier, I., Desrochers, A. and Campbell, K. (2000). Multiple P3 evidence of a two-stage process in word gender decision. *Cognitive Neuroscience and Neuro-psychology*, 11, 3527–3531.

New, B., Pallier, C., Ferrand, L. and Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique*, 101: 447–462.

Pergnier, M. (1989). *Les anglicismes. Danger ou enrichissement pour la langue française?* Paris: Presses Universitaires de France.

Poirier, J. (2012). Pronoms clitiques objects: fréquence et expression du genre grammatical dans le discours oral des enseignants en classes d'immersion. MA dissertation, Université du Québec à Montréal.

Poirier, J. and Lyster, R. (2014). Les pronoms objets directs de la 3e personne et leur apport aux indices du genre grammatical dans le discours oral des enseignants en immersion. *La Revue canadienne des langues vivantes*, 70: 246–267.

Presson, N., MacWhinney, B. and Tokowicz, N. (2014). Learning lexical means: the use of rules by novice learners. *Applied Psycholinguistics*, 35: 712–713.

Price, G. (2008). *A Comprehensive French Grammar*. Oxford /Malden, MA: Blackwell.

Roché, M. (1992). Le masculin est-il plus productif que le féminin? *Langue Française*, 96: 113–124.

Sabourin, L. (2001). L1 effects on the processing of grammatical gender in L2. In: S. Foster-Cohen and A. Nizegorodcew (eds.), *EUROSLA Yearbook*. Amsterdam: John Benjamins, pp. 159–169.

Sabourin, L., Stowe, L. A. and Haan, G. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22: 1–29.

Sagarra, N. and Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, 120: 2022–2039.

Sagarra, N. and Herschensohn, J. (2012). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, 17: 1–21.

Savary, A. (2000). Recensement et description des mots composés. Méthodes et applications. PhD thesis, Université Paris 7.

Schafroth, E. (2003). Gender in French: structural properties, incongruences and asymmetries. In: M. Hellinger and H. Bußmann (eds.), *Gender Across Languages*. Amsterdam: John Benjamins, pp. 87–117.

Séguin, H. (1969). Les marques du genre dans le lexique français écrit contemporain: compilation des cas et essai de classement. PhD thesis, Université de Montréal.

Seigneuric, A., Zagar, D., Meunier, F. and Spinelli, E. (2007). The relation between language and cognition in 3- to 9-year-olds: the acquisition of grammatical gender in French. *Journal of Experimental Child Psychology*, 96: 229–246.

Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodriguez, W. and Dubois, D. P. (2002). When language affects cognition and when it does not: an analysis of grammatical gender and classification. *Journal of Experimental Psychology*, 131: 377–397.

Shi, R. and Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, 15: 517–533.

Spalek, K., Franck, J., Schriefers, H. and Frauenfelder, U. (2008). Phonological regularities and grammatical gender retrieval in spoken word recognition and word production. *Journal of Psycholinguistic Research*, 37: 419–442.

Spilka, I. (1976). Assessment of second language performance in immersion programs. *The Canadian Modern Language Review*, 32: 543–557.

Spinelli, E. and Alario, F.-X. (2002). Gender context effects on homophone words. *Language and Cognitive Processes*, 17: 457–469.

Surridge, M. (1986). Genre grammatical et dérivation lexicale en français. *Canadian Journal of Linguistics*, 31: 267–284.

Surridge, M. (1993). Gender assignments in French: The hierarchy of rules and the chronology of acquisition. *International Review of Applied Linguistics*, 31: 77–95.

Swain, M. (2000). French immersion research in Canada: recent contributions in SLA and applied linguistics. *Annual Review of Applied Linguistic*, 20: 199–212.

Taft, M. and Meunier, F. (1998). Lexical representation of gender: a quasi-regular domain. *Journal of Psycholinguistic Research*, 27: 23–45.

Tarone, E., Frauenfelder, U. and Selinker, L. (1976). Systematicity/variability and stability/instability in interlanguage systems: more data from Toronto French immersion. *Papers in Second Language Acquisition*, special issue of *Language Learning*: 93–134.

Taylor-Browne, K. (1984). The acquisition of grammatical gender by children in French immersion programs. MA dissertation, University of Calgary, Canada.

Trudgill, P. (1999). Language contact and the function of linguistic gender. *Poznan Studies in Contemporary Linguistics*, 35: 133–152.

Tucker, G. R. (1968). French speakers' skill with grammatical gender: an example of rule-governed behavior. PhD thesis, McGill University, Montreal.

Tucker, G., Lambert, W., Rigault, A. and Segalowitz, N. (1968). A psychological investigation of French native speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7: 312–316.

Tucker, G., Lambert, W. and Rigault, A. (1977). *The French Speaker's Skill with Grammatical Gender: An Example of Rule-governed Behavior*. Paris: Mouton.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, 92: 231–270.

van der Linden, E. and Hulk, A. (2009). Having a hard time: acquiring gender in 2L1/child L2. Talk given at the EUROSLA Conference, Cork, Ireland.

Violin-Wigent, A. (2006). Gender assignment to nouns code-switched to French: observations and explanations. *International Journal of Bilingualism*, 3: 253–276.

Wagner, R. L. and Pinchon, J. (1991). *Grammaire du français classique et moderne*. Paris: Hachette.

Warden, M. (1997). The effect of form-focused instruction on control over grammatical gender by French immersion students in grade 11. PhD thesis, University of Toronto.

APPENDIX A — STIMULI BY SESSION

Session 1	Session 2	Session 3
1. uvule	hymne	termite
2. camion-citerne	greffe	effluve
3. compatriote	casse-cul	tête-à-tête
4. artiste	casse-noisettes	vermine
5. carpe	poêle	éloge
6. élève	tour	idole
7. cache-flamme	papier-toilette	libraire
8. modéliste	photographe	exode
9. collègue	guide	portefeuille
10. bugle	mémoire	aéronef
11. porte-parole	mort	holocauste
12. dentiste	garde-cendre	antipode
13. domestique	poste	coupe-gorge
14. avant-guerre	faune	vedette
15. pongiste	gîte	Pâque
16. adulte	mousse	bateau-mouche
17. aromatisation	aide	météore
18. adversaire	interligne	antidote
19. mime	chou-fleur	nourrisson
20. complice	ponte	porte-cigarettes
21. entrejambe	pendule	ovale
22. bibliothécaire	trompette	amour
23. brise-glace	coche	crypte
24. philosophe	garde-pêche	médecin
25. sentinelle	vague	pastiche
26. pianiste	timbre-poste	bateau-citerne
27. pantomime	nomade	sage
28. reporter	page	auspices
29. superbe	oiseau-mouche	recrue
30. homme-grenouille	diesel	bouche-à-bouche
31. secrétaire	face-à-main	mappemonde
32. épitaphe	vigile	parapluie
33. convive	physique	célibataire
34. ayant cause	micro-cravate	opuscule
35. fantaisiste	ange	cache-prise
36. pupille	pause-café	sous-main
37. internaute	cartouche	scarabée
38. chimiste	ascendant	primeur
39. jeune	platine	lâche
40. universitaire	face-à-face	poulpe
41. malade	cancre	chèvrefeuille
42. primeurs	radio	porte-voix
43. ministre	taupe	astérisque
44. partenaire	pensionnaire	squelette
45. mi-temps	faire-part	porte-plume
46. esclave	garde	arministice

145

47. patriote	casse-cou	équivoque
48. basque	chose	tribut
49. orange	moule	autoradio
50. camarade	chef	chenapan
51. boum	abîme	violoncelle
52. après-guerre	gourmet	dédale
53. membre	légume	croque-madame
54. casse	fonctionnaire	apogée
55. comique	bloc-note	jeûne
56. crêpe	enseigne	garde-boue
57. môme	maire	proches
58. barde	essuie-mains	circulaire
59. gosse	nouveau-né	aérogare
60. espace	critique	porte-jarretelles
61. en-tête	garde-malade	annexe
62. foudre	prodige	variable
63. renoncule	interprète	mécène
64. pique-assiette	cache-misère	alcôve
65. manche	filou	casse-gueule
66. crème	pied-à-terre	flegme
67. garde-à-vue	adobe	Pâques
68. touriste	porte-à-faux	fécule
69. merci	canaille	crève-la-faim
70. lustucru	fantôme	mausolée
71. mode	croque-mort	libido
72. parachute	fripouille	intervalle
73. finale	potache	parapente
74. manœuvre	tribu	ombre
75. plastique	aéronaute	appui-tête
76. politique	après-midi	lutin
77. haut-de-forme	carpette	sans-gêne
78. geste	pétale	personne
79. solde	victime	obélisque
80. tête-à-queue	journaliste	porte-monnaie
81. livre	cache-poussière	bébé
82. parallèle	anathème	trophée
83. ouvre-boîte	ministre	mitaine
84. vase	nouille	idylle
85. soutien-gorge	tiers	coupe-faim
86. somme	pare-brise	œuvres
87. argile	arome	amours
88. goutte-à-goutte	sentinelle	brûle-gueule
89. œuvre	jarretelle	teigne
90. porte-bouteilles	juge	propriétaire
91. cave	bouche-bouteille	salamandre
92. chèvre	monstre	orbite
93. délices	chaume	voiture-restaurant
94. minuit	croque-mitaine	cache
95. génie	balance	casse-croûte
96.	espèce	cul-de-poule

146

97.	teigne	épiderme
98.	vapeur	victuailles
99.		oasis
100.		stalactite
101.		gens
102		connaissance
103		dupe

APPENDIX B

Strategies participants used when they were unsure of the gender of a noun. The original French comments were translated (translation mine).

– "*J'ai utilisé un participe passé avec lequel l'accord féminin est phonétiquement réalisé*" ('I used a past participle with a phonetically realized feminine agreement').

– "*J'utilise la connaissance du monde. Par exemple pour 'proches', il est évident que le mot va servir pour des hommes et des femmes, d'où M et F*" ('I use world knowledge. For example, for *proches* 'loved ones', the word obviously refers to both men and women, hence M and F').

– "*J'ai essayé de mettre le mot dans un contexte significatif. Par exemple pour le mot 'Pâques', je n'ai pas pu trouver une phrase qui permette de définir si le mot est masculin ou féminin: 'la fête de Pâques' ou 'nous sommes allés à Pâques à . . . ', alors que la Pâque est pour moi nettement féminin si on la considère comme un nom*" ('I tried to place the word in meaningful context. For example, for the word *Pâques* 'Easter', I could not find a sentence that could define the word as masculine or feminine: 'the-FEM Easter holiday' or 'at Easter, we went to . . . ', but for me Easter is clearly feminine if it is considered as a noun').

– "*J'ai construit une phrase*". ('I created a sentence').

– "*Je disais directement le mot dans ma tête avec 'un', 'une', 'le', 'la'*". 'I would say the word mentally with 'a'–MSC, 'a–FEM', 'the–MSC', 'the–FEM').

– "*J'ai utilisé des mots composés similaires; ou vu si le premier mot dans les noms composés étaient des verbes ou des noms… parfois c'était aussi "le souvenir" dans ex ci-dessous je cherche le sujet clef! autre: agonie avec combien de… j'ai souvent changé d'avis MAIS probablement parce que j'enseigne…. les fleurs? J'en ai acheté (on ne fait pas l'accord avec 'en' si je souviens bien) mais 'combien en as-tu achetées' je fais l'accord instinctivement… (au son)*". ('I used similar compound words, or tried to see if the first word in the compound words were verbs or nouns..sometimes, it was what I remembered, in the example below, I'm looking for the subject. Other: I agonized with 'how many' . . . I often changed my mind BUT that's probably because I'm a teacher . . . flowers? I bought some (there is no agreement with 'some' if I remember correctly), but in 'how many did you buy?', I do the agreement instinctively (based on the sound)'.

– "*J'ai essayé avec les deux et j'ai choisi celui qui m'avait l'air le plus naturel*". ('I tried with both and picked the one that seemed the most natural').

– "*Pour les mots composés dont je ne connaissais pas le sens, j'ai procédé par analogie*". (''when I didn't know the meaning of compound words, I proceeded by analogy'.

147

– "*Quand je ne savais vraiment pas, j'ai essayé de décomposer le mot et de faire des hypothèses sur son étymologie*". ('When I really didn't know, I tried to decompose the word and to guess its etymology').

– "*J'ai utilisé les mots avec différents adjectifs comme beau/belle, bon/bonne mais aussi avec des articles pour être vraiment certain, mais je suis sûr de m'être quand même trompé*". (I used words with different adjectives such as beautiful-MSC/FEM, good –MASC/FEM, but also with articles to be really sure, but I'm certain I still made mistakes').

148