The bilingual mental lexicon in L2 sentence processing

Holger Hopp
Technische Universität Braunschweig

Abstract
This paper explores the consequences of the language-integrated nature of the bilingual mental lexicon for L2 sentence processing. It reviews L2 processing studies on gender agreement and syntactic structure building that test whether delays and cross-linguistic influence in lexical processing of the L2 lead to differences between L2 and L1 sentence processing. Slower L2 lexical processing delays and attenuates effects of syntactic structure in L2 sentence processing. In addition, cross-linguistic lexical influence can engender non-target patterns in L2 compared to L1 sentence processing. The paper spells out the assumptions and predictions of the Lexical Bottleneck Hypothesis, and I discuss how its insights can be incorporated into current models of L2 sentence processing.

1. Introduction

Research on L2 sentence processing uses on-line methods to study how L2 learners incrementally build syntactic structure and derive interpretation in real-time comprehension. Studying the dynamics of the human processing mechanism, i.e. the parser, in L2 comprehension gives insights into whether the parser employs comparable processing strategies as in the L1 and whether L2 users recruit the same types of grammatical and non-linguistic knowledge in sentence comprehension as monolinguals do (for review, see Roberts, 2013). Beyond revealing the dynamics of sentence processing in an L2, L2 processing
research holds the promise of elucidating whether the persistent problems adult L2 learners have in mastering L2 morphosyntax are in part caused or magnified by difficulties in processing L2 grammatical information in a target-like manner.

Several approaches to L2 sentence processing indeed propose that non-target sentence comprehension among L2 learners follows from differences between L2 learners and monolinguals in parsing. For instance, L2 learners may overrely on non-syntactic information, e.g. lexical, discourse or frequency information in parsing (e.g. the Shallow Structure Hypothesis, Clahsen & Felser, 2006, 2017; Cunnings, 2017). Other types of models, e.g. Sorace’s Interface Hypothesis (Sorace, 2011) and capacity models of L2 processing (McDonald, 2006; Hopp, 2010) claim that bilinguals fail to integrate multiple information types as efficiently as monolinguals do in real-time sentence processing.

The present paper adopts a different perspective on non-target L2 sentence processing. It explores the possibility that some aspects of non-target L2 parsing may not follow from any differences between L2 and L1 sentence processing; rather, they may be the consequence of the architecture of the processing system, in particular, the characteristics of the bilingual mental lexicon. Since lexical processing partly precedes and feeds into syntactic processing, key characteristics of bilingual lexical processing may cause aspects of non-target parsing that have so far been interpreted as signatures of qualitative differences between L1 and L2 sentence processing.

Research on bilingual word recognition and production found that bilingual lexical representations are interconnected across languages and accessed non-language selectively (for review, see Tokowicz, 2015). In this paper, I am going to explore the consequences of these aspects of the bilingual mental lexicon for L2 sentence processing. Specifically, I will give a selective overview of some recent studies on L2 sentence processing that have addressed relations between lexical processing and L2 sentence processing. I will suggest that
explicit attention to lexical processing needs to be part and parcel of any approach to L2 sentence processing.

In my review, I will focus on two areas of sentence processing in which pronounced differences have been reported between adult L2 and monolingual processing, i.e. grammatical gender and building hierarchical structural relations. For each area, I discuss studies that have systematically considered how bilingual lexical processing gives rise to non-target sentence processing. The paper concludes with a general consideration of the role of lexical processing in sentence comprehension, it presents the Lexical Bottleneck Hypothesis and sketches how this hypothesis can be integrated into current models of L2 sentence processing.

2. The Bilingual mental lexicon

Experimental investigations of the bilingual mental lexicon marked the beginning of psycholinguistic research on bilingual language processing (for review, see Kroll & Tokowicz 2005). Initial research address the issue whether bilinguals had two separate or one shared lexicon(s) for each of their languages by comparing the recognition and production of language-particular versus language-ambiguous word forms, e.g. cognates or interlingual homographs. Subsequently, attention shifted to studying the dynamics of lexical retrieval or activation. Research on bilingual word recognition and production uncovered two major aspects in which the dynamics of L1 and L2 lexical processing differ: (A) Lexical retrieval is slower and is associated with larger frequency effects in L2 than in L1 processing; (b) lexical access is non-language-selective in that bilinguals activate lexical representations across languages in production and comprehension.

Evidence of slower lexical retrieval comes from the comprehension and production of language-particular word forms, i.e. L2-only words, in bilinguals. For instance, in lexical
decision tasks, L2 speakers react more slowly than monolingual speakers, and they make
more decision errors (e.g. Lemhöfer, Dijkstra, Schriefers, Baayen, Grainger, & Zwitserlood,
2008). Similarly, in picture naming, L2 speakers show a larger number of retrieval errors, e.g.,
false starts, tip-of-the-tongue effects and slips of the tongue (Kroll & Hermans, 2011). Slower
lexical retrieval among L2 learners also obtains in sentence contexts, i.e. when they produce
or read sentences (for production, e.g., Runnqvist, Gollan, Costa, & Ferreira, 2013; for
reading, Gollan, Slattery, Goldenberg, van Assche, Duyck, & Rayner, 2011). On top of being
generally slower than monolinguals, L2 speakers also demonstrate larger frequency effects in
that L2 readers slow down disproportionately more on low-frequency words compared to
monolingual readers. The Weaker Links hypothesis (Gollan, Montoya, Cera, & Sandoval,
2008) or Frequency-lag hypothesis (Gollan et al., 2011; see also Kroll & Gollan, 2014)
interprets slower lexical retrieval in bilinguals as a direct consequence of the lower frequency
with which bilinguals access word forms in each language (see also Segalowitz & Segalowitz,
1993). Since bilinguals speak and understand each of their languages relatively less often than
monolinguals, they activate word forms less often than monolinguals do, which, in turn,
results in slower lexical retrieval. The effects of weaker lexical links are magnified for low-
frequency items. The frequency lag in bilingual word recognition correlates with language
proficiency reflecting exposure to and use of the L2 (Diependaele, Lemhöfer, & Brysbaert,
2013). Indeed, the size of frequency effects of word recognition in bilingual reading relates to
relative use of the respective language, irrespective of whether bilinguals read in the L1 or the
L2 (Whitford & Titone, 2012). In conjunction, these findings show that links between word
forms, lemma and conceptual representations are strengthened through use which leads to
greater automatization in lexical processing.

Evidence of non-selective lexical access comes from research on word forms that are
ambiguous between languages, i.e. cognates like English–German film. Even when a task is in
the L1 or L2 only, L2 speakers, yet not monolinguals, process cognate words faster compared
to lexically matched language-particular word forms (e.g., *movie*). By contrast, interlingual homographs (false friends) which share form, yet not meaning across languages are associated with slowdowns (e.g., Libben & Titone, 2009). Cognate facilitation and homograph inhibition indicate that lexical representations are activated across all languages of a bilingual even when only one language is being used. Current models of the bilingual mental lexicon, e.g. the Bilingual Interactive Activation Plus (BIA+) model (Dijkstra & van Heuven, 2002), accommodate these effects by postulating that there is a single integrated lexicon and that context, e.g. the language of the task, does not affect early stages of word recognition. Indeed, when cognates or interlingual homographs are embedded in sentences presented exclusively in the L1 (e.g., Van Assche, Duyck, Hartsuiker & Diependaele, 2009; Titone, Libben, Mercier, Whitford & Pivneva, 2011) or the L2 (e.g., Schwartz & Kroll, 2006; Duyck, Van Assche, Drieghe & Hartsuiker, 2007; Libben & Titone, 2009, Van Assche, Drieghe, Duyck, Welvaert & Hartsuiker, 2011), bilinguals continue to show cognate facilitation or homograph inhibition effects. Hence, cross-linguistic lexical activation perseveres even if the sentence context signals that the other language is irrelevant. The strength of cognate effects is affected by proficiency. More advanced L2 learners evince less cognate facilitation (e.g. Libben & Titone, 2009), yet even highly proficient L2 readers continue to process cognates faster than control words in the L1 and the L2 (see Titone et al., 2011; van Hell & Tanner, 2012). In conjunction, these findings of differential processing of cognates and interlingual homographs compared to language-particular words indicate that lexical representations are activated non-language selectively across proficiency stages of bilingualism even when bilinguals read or process in a single-language context. In sum, research on bilingual word recognition demonstrates that weaker links lead to slower lexical processing and larger frequency effects in the L2 and that cross-linguistic activation leads to different patterns of lexical activation in bilingual compared to monolingual processing.
In the real-time comprehension of sentences, aspects of lexical processing subserve parsing, since lexical information is relevant for agreement processing and structure building. Models of (monolingual) sentence processing differ in the status they ascribe to lexical information in parsing. They range from the Garden-Path model (e.g. Frazier, 1987), in which lexical information in parsing is initially largely limited to word class, to lexicalist constraint-based models (e.g. MacDonald, Pearlmanual, & Seidenberg, 1994), according to which all aspects of lexical information guide and inform parsing.

Under any account, lexical processing in bilinguals may impact parsing in quantitative and qualitative ways. First, lexical processing might cause additive delays in that, e.g., slower lexical retrieval will cause slower syntactic processing. Second, lexical processing might interact with parsing in that slower lexical retrieval and cross-linguistic lexical activation cause different syntactic processing patterns.

In the following sections, I summarize a growing body of research that addresses the consequences of lexical processing in the bilingual mental lexicon for L2 sentence processing. I focus on sentence comprehension, although work on sentence production, especially priming studies in the framework of the Shared Syntax Model (Hartsuiker & Pickering, 2008), also assess lexical and syntactic cross-linguistic interactions. In both primed and unprimed sentence production, lexical overlap and L1 lexical frequency have been found to affect L2 sentence production (for review, Jackson, Massaro, & Hopp, 2017). I first turn to agreement processing of grammatical gender and discuss effects of weaker links and non-selective lexical access. Then, I discuss these effects in the context of grammatical structure building.

3. Grammatical gender

Grammatical gender lends itself particularly well for studying possible interactions between lexical and grammatical processing, since grammatical gender comprises lexical and syntactic
aspects. For one thing, learners need to assign the target lexical gender to a noun when they learn it, e.g. they need to classify the Spanish noun *casa* as feminine (*casaFEM*). Gender assignment encompasses the allocation of a noun to a set of feminine nouns and the linking of a noun to a gender node at the lemma level (e.g., Vigliocco, Antonini, & Garrett 1997). Second, learners need to compute gender agreement, i.e. they realize gender inflection on dependents of the noun, e.g. determiners, adjectives or pronouns (*la casa roja* – *theFEM houseFEM redFEM*). Gender agreement is a syntactic process which proceeds by feature checking or matching processes across different constituents (e.g., Carstens 2000). In real-time comprehension, learners accordingly need to access the lexical gender of a noun or a dependent and then match it with other agreeing constituents (e.g. Franck et al., 2008). Both aspects of gender present problems to adult L2 learners, especially if the L1 does not realize gender.

Even at advanced to near-native proficiency levels, L1 English learners of Spanish or German achieve only between 75% to 90% accuracy in gender assignment for nouns in picture naming or description tasks (e.g. Alarcón, 2011; Bruhn de Garavito & White, 2003; Franceschina, 2005; Grüter, Lew-Williams, & Fernald, 2012; Hopp, 2013). These findings suggest that gender assignment remains variable in an L2, with learners having non-target or unstable gender representations for some nouns in their lexicons.

Adult L2 learners also show difficulties in computing gender agreement. In reaction-time or neurophysiological tasks, learners display varying sensitivity to gender agreement violations between nouns and determiners or adjectives (e.g. *la casa roja* – *theFEM houseFEM redMASC*), in particular if the L1 does not use gender agreement between nouns and its dependents (e.g. Sabourin & Stowe, 2008; though see Foucart & Frenck-Mestre, 2011). Problems with gender agreement are aggravated in predictive processing, which is often studied in visual-world eye tracking. As participants listen to sentences containing gender marked articles, they look at a display showing various objects. Recordings of eye movements
reveal the extent to which they use gender marking on articles to make anticipatory looks to referents during sentence comprehension (for review, Roberts & Siyanova-Chanturia, 2013). Adult L2 learners routinely fail to use gender marking on a determiner to anticipate an upcoming noun (e.g. \( \text{la}_{\text{FEM}} \rightarrow \text{casa}_{\text{FEM}} \); Grüter et al., 2012; Lew-Williams & Fernald, 2010; though see Dussias et al., 2013).

So far, researcher have treated variability in gender assignment and non-target gender agreement as unrelated, since the construction of a structural agreement relationship is, in principle, independent of the target assignment of lexical gender. In fact, several studies report that L2 learners evince consistent gender agreement despite non-target gender assignment (e.g. \( \text{el casa rojo} \rightarrow \text{the}_{\text{MASC}} \text{house}_{\text{MASC}} \text{red}_{\text{MASC}} \); e.g. White et al., 2004). Similarly, Lemhöfer et al. (2014) show in an ERP study that learners are sensitive to violations of their subjective gender assignment, i.e. they compute gender agreement violations against their own (sometimes non-target) gender assignment values. However, evidence that gender assignment and agreement errors are unrelated comes mainly from tasks measuring sensitivity to gender agreement violations. Recent studies suggest that lexical variability in gender marking becomes relevant in predictive processing of gender agreement.

3.1. Weaker Links and predictive gender processing

In the context of predictive processing, a study by Hopp (2013) set out to explore potential relations between gender assignment and agreement. It builds on the reasoning that non-target lexical gender assignment leads to unsuccessful prediction in processing. If a learner assigns a non-target gender to a noun (e.g., \( \text{casa}_{\text{MASC}} \)), she will make an erroneous prediction about the upcoming noun (\( \text{el}_{\text{MASC}} \rightarrow \text{casa}_{\text{MASC}} \)), which will immediately be flagged as incorrect by the input when a different noun follows the article (\( \text{el}_{\text{MASC}} \ldots \text{libro}_{\text{MASC}} \)). The resulting prediction error serves as a signal to the listener to revise their prediction. In parsing, listeners have
been found to adjust or attenuate their predictions following prediction errors (e.g. Fine & Jaeger, 2013; DeLong, Troyer, & Kutas, 2014). Finding that listeners adjust their predictions underscores the adaptivity of the parser that seeks to maximize the utility of prediction in processing (for review, Kuperberg & Jaeger, 2016). Hopp argued that lexical variability in gender assignment curtails the utility of using gender for prediction, since non-target lexical gender leads to prediction error. In consequence, learners will attenuate the use of gender for prediction. Conversely, learners who have acquired target lexical gender assignment should use gender for prediction, since gender proves to be a successful prediction cue. Consistent with this account, a group of highly proficient L1 English learners of German who had target gender assignment in elicited production performed indistinguishably from native Germans in predictive processing (Hopp, 2013). In contrast, a group of L1 English learners with partially variable L2 gender assignment did not use gender marking for prediction in comprehension, even if only the subset of nouns was considered for which they showed target gender assignment in production. Hence, target lexical gender assignment and predictive gender agreement processing correlate.

In a training study, Hopp (2016) further considered causal relations between gender assignment and agreement processing. In a pre-post test training study, 34 L1 English learners of German were first tested on their lexical gender assignment and their predictive processing. A week later, they received training on gender assignment of the nouns used in the experiment, and they subsequently performed the predictive processing task again. In the pre-test, the group did not make use of gender agreement in predictive processing, while they did in the post-test. The size of the prediction effect was strongly correlated with the accuracy of target gender assignment in the post-test. Only the learners who had converged on target gender assignment of all nouns in the experiment could use gender marking on articles for predictive agreement processing. Hence, the study shows that target gender assignment is a prerequisite for target gender agreement processing. If lexical gender assignment remains
variable, learners do not use gender in agreement processing. To test whether lexical variability would disrupt predictive gender agreement processing more generally, a second experiment on native German speakers interspersed lexical gender mistakes in the filler items. By introducing lexical variability in the input to L1 speakers, the experiment aimed to emulate the lexical variability typical of the lexicons of L2 learners. Once gender mistakes had accrued in the filler items over the course of the experiment, native German speakers stopped using gender predictively for the experimental items (see also Hanfluková, van Alphen, van Goch, & Weber, 2012). In other words, lexical variability in the input to L1 speakers affected gender prediction in ways similar to how lexical variability in the L2 lexicon constrains predictive gender agreement processing.

In conjunction, these findings bolster the contention that relations between lexical and grammatical aspects of gender are not merely correlational but causal in adult L2 processing. Importantly, the effects of lexical variability go beyond accuracy in gender processing of specific lexical items. Instead, weak and unstable lexical gender representations have consequences for gender agreement processing of all nouns. Once the parser experiences prediction errors with some nouns, the parser adjusts prediction according to gender generally. These results from gender processing chime with studies finding that the reliability of a cue to allow target prediction determines its use in processing (Henry, Hopp, & Jackson, 2017).

It is currently an open question whether these effects extend beyond predictive processing, in which accuracy of lexical gender assignment is a prerequisite for target predictive agreement processing. Recently, a study on gender violations using ERPs also reported relations between lexical variability in production and agreement in comprehension (Alemán Bañón, Miller, & Rothman, 2017). Further, there may be different sources of variability in gender assignment. Lexical variability may reflect weaker links in retrieval of gender information from the lexicon, since the retrieval of gender information in an L2 is less
robust than in an L1 (Shantz & Tanner, 2016). In addition, lexical variability may be due to a weaker or unstable representation of gender in the adult L2 lexicon, because L2 learners learn gender in an L2 differently from how children acquire gender in an L1 (see the Lexical Gender Learning Hypothesis; Hopp 2013, building on Grüter et al., 2012; Arnon & Ramscar, 2012). Current research explores these aspects in training studies on artificial languages (Shantz & Tanner, 2017).

3.2. Non-selective lexical access and predictive gender processing

L2 learners whose L1 marks grammatical gender have more target gender agreement than learners whose L1s lack gender in off-line and on-line tasks (e.g. White et al., 2004; Sabourin & Stowe, 2008; though see Loerts, 2011). However, among learners whose L1 also realizes gender agreement processing is not invariably native-like, since target predictive gender agreement processing can be limited to certain gender markings. For instance, Dussias et al. (2013) found that highly proficient English learners of Spanish, yet not less proficient learners, used gender on articles as a predictive agreement cue, while (low-proficient) Italian learners of Spanish showed predictive gender processing only for feminine, yet not masculine articles. Given that Italian, yet not English, has grammatical gender, this asymmetry suggests that cross-linguistic differences in gender marking affect L2 gender processing.

In research on the bilingual mental lexicon, such differences have been studied in terms of gender congruency (Paolieri, Cubelli, Macizo, Bajo, Lotto, & Job, 2010). Lexical gender congruency refers to similarities in gender assignment between a noun and its translation equivalent. For instance, the Spanish noun *vela* (candleFEM) has the same gender as the noun *Kerze* (candleFEM) in German, while *maleta* (suitcaseFEM) is gender-incongruent with the German *Koffer* (suitcaseMASC). A substantial body of research on word recognition and picture naming found that bilingual speakers are faster and more accurate in recognizing and
producing gender-congruent than gender-incongruent words (e.g. Klassen, 2016). To account for these findings, models assume that gender representations in the bilingual mental lexicon are interrelated or accessed language non-selectively (e.g. Salamoura & Williams 2007). Gender congruency also affects retrieval of gender in sentence contexts in that, e.g., Italian-Spanish bilinguals show greater predictive processing of target objects if they are gender congruent (Morales, Paolieri, Dussias, Valdés Kroff, & Gerfen, 2015; see also Weber & Paris, 2004). In a series of studies (Hopp & Lemmerth, 2017; Lemmerth & Hopp, in press), we probed the extent to which gender congruency would constrain predictive gender agreement processing in German by Russian-German bilinguals.

Both German and Russian assign nouns to one of three gender classes, i.e. masculine, feminine and neuter, and gender agreement is realized inside noun phrases. German marks gender overtly on articles (1a), while Russian – for lack of articles – marks gender on noun endings (1b).

(1) a. der\textsubscript{MASC}/die\textsubscript{FEM}/das\textsubscript{NEUT} Tisch\textsubscript{MASC}/Lampe\textsubscript{FEM}/Kleid\textsubscript{NEUT} (German)
   the table/lamp/dress
b. \textnothing stol\textsubscript{MASC}/lamp\textsubscript{FEM}/plat\textsubscript{NEUT} (Russian)
   table/lamp/dress

(2) a. (ein) roter\textsubscript{MASC}/-es\textsubscript{NEUT} Tisch\textsubscript{MASC}/Kleid\textsubscript{NEUT} (German)
   a red table/dress
b. \textnothing krasn-i\textsubscript{MASC}/-o\textsubscript{NEUT} stol/plat\textsubscript{je} (Russian)
   red table/dress

Both languages have prenominal attributive adjectives that mark gender on endings (2). In an experiment that crossed the factors gender congruency and syntactic marking of gender, Hopp & Lemmerth (2017) tested whether 24 L1 Russian intermediate to advanced adult L2 learners
of German would use gender marking in predictive processing. The results showed an interaction of congruency and proficiency. Advanced learners used German gender agreement predictively, irrespective of gender congruency and syntactic marking. Intermediate learners equally demonstrated predictive gender agreement processing when gender marking was realized on adjectives, yet they could use gender realized on articles for agreement only if the nouns were gender-congruent in Russian. In other words, lexical overlap in gender facilitated gender agreement processing (for comparable findings from bilingual Russian-German children, see Lemmerth & Hopp, in press). Critically, lexical congruency effects only held in syntactic contexts in which Russian and German differ, i.e. gender marking on articles. In gender processing, then, lexical and syntactic factors interact in that lexical overlap can facilitate gender agreement processing in syntactic contexts that are dissimilar in L1 and L2.

In sum, research on L2 gender processing suggests that both weaker lexical links and non-selective lexical access have consequences for sentence processing in that agreement processing is constrained by lexical factors, especially in syntactic contexts where L1 and L2 differ. Importantly, these effects are paradigmatic. First, instability in gender assignment for some nouns has implications for gender processing generally. Lexical variability with some nouns appears to attenuate predictive gender agreement processing for all nouns. Second, incongruency in gender assignment for some nouns has implications for gender processing in that predictive use of gender is limited to congruent lexical realizations of gender in syntactic contexts that differ between the L1 and the L2.

Finally, the studies summarized above point to interactions between lexical and syntactic processing. Lexical and syntactic processing interact in that differences in the representation and processing of lexical gender lead to qualitatively different patterns of syntactic agreement processing in L2 learners and monolinguals. However, one may argue that these interactions are perhaps particular to grammatical gender for which lexical and
syntactic aspects are closely intertwined. In the following section, I discuss a selection of studies that consider effects of lexical processing for syntactic structure building.

4. Building grammatical structure

Two broad lines of research have studied how L2 learners build grammatical structure in the real-time comprehension of ambiguous or syntactically complex sentences. The first line of research probes which information types guide initial L2 parsing and whether L2 learners can successfully reanalyse a parse provided the initial parsing preference has led to an incorrect interpretation (e.g. Frenck-Mestre & Pynte, 1997; Hopp, 2014a; Jackson, 2008; Jacob & Felser, 2016, Jegerski, 2012; Roberts & Felser, 2011). Many studies investigated structural parsing preferences in ambiguous relative clauses, especially since parsing preferences vary cross-linguistically (e.g. Felser, Roberts, Marinis, & Gross, 2003; Papadopoulou & Clahsen, 2003, Rah & Adone, 2010). A second line of research targets abstract syntactic structure in filler-gap dependencies, e.g. traces, or syntactic locality restrictions, e.g. Subjacency or Binding Principles. For both lines of research, studies report that L2 parses are guided by plausibility and discourse information (e.g. Hopp, 2014a; Pan, Schimke, & Felser, 2015; Roberts & Felser, 2011) while robust reliance on syntax is less in evidence (e.g. Felser et al., 2003; Marinis, Roberts, Felser, & Clahsen, 2005; though see Omaki & Schulz, 2011).

4.1. Weaker Links and syntactic structure building

Several studies have examined effects of lexical retrieval speed of the words contained in the syntactic structures that L2 learners build incrementally. One type of studies has considered individual differences between L2 learners in lexical decoding facility (e.g. McDonald & Roussell, 2010). In a study on relative clause attachment, Hopp (2014b) tested 75 high-
intermediate to advanced L1 German learners of English. They took a reading-span task to assess working memory, a lexical decision task to probe lexical decoding, and they read temporarily ambiguous relative clauses that were either disambiguated locally by subject-verb agreement (3a) or non-locally by gender marking on reflexive pronouns (3b).

(3)  
  a. The director congratulated the instructor of the schoolboys who was writing the reports.
  b. The student had liked the secretary of the professor who had almost killed himself in the office.

For the locally disambiguated sentences, the L2 group demonstrated a clear high attachment preference in eye-movement measures. For the non-locally disambiguated sentences, lexical decoding, yet not working memory, interacted with reading measures. Only L2 readers with fast lexical decoding skills had a native-like structural attachment preference. These findings underscore that structure building in the creation of non-local dependencies depends in part on efficient lexical decoding of the words implicated in the parse.

Further evidence of the importance of lexical aspects comes from processing studies that manipulate lexical retrieval speed via word frequency. Building on a study by Tily, Fedorenko and Gibson (2010), Hopp (2016b) considered differences in the processing of subject versus object cleft sentences as in (4).

(4)  
  a. It was Amanda who, it scared/horrified Sulena with a frightening look. (subject cleft)
  b. It was Amanda who, Sulena scared/horrified it with a frightening look. (object cleft)
Object clefts give rise to processing delays that reflect the differences in syntactic structure between subject and object clefts (e.g. Clifton & Frazier, 1989). The verb in the sentence was either a high-frequency verb (scare) or a semantically equivalent lower-frequency verb (horrify). Reading times of the underlined segments in (4) showed a significant four-way interaction of syntactic structure, verb frequency, segment and group. Unlike the L1 group, the L2 learners were strongly affected by the lexical frequency of the verb in that reading slowdowns associated with the more complex object cleft structure showed in different segments. For high-frequency verbs, L2 readers evinced reading delays for object clefts in the cleft segment (scared Sulena/Sulena scared). For low-frequency verbs, these slowdowns arose only in the postcleft segment (with a). Hence, slower lexical retrieval with low-frequency verbs delays effects of syntactic structure. These results point to a functional staging of lexical and syntactic processing: structure-building operations are restricted to be implemented only after lexical retrieval of the items they incorporate has been completed to a certain extent (see also Tily et al., 2010; Staub, 2011). Since L2 learners are slower in lexical decoding and have less robust representations of low-frequency items (Gollan et al., 2011), frequency effects occur sooner and in a more pronounced fashion in L2 than in L1 processing. Other studies report that target-like L2 processing is modulated by reading speed (e.g. Kaan, Ballantyne; & Wijnen, 2014; Roberts & Felser, 2011). Since overall reading speed subsumes lexical processing, these studies also support the idea that processing at levels other than syntactic structure building contributes to target-like syntactic processing.

In conjunction, these studies show that the slower timing of lexical retrieval in L2 speakers affects sentence processing. Both lower lexical decoding ability and stronger lexical frequency effects implicate a slower time-course of lexical processing, which, in turn, may entail that structure building is delayed, attenuated or even absent.

4.2. Non-selective lexical access and syntactic structure building
In L2 processing, L2 learners access lexical-thematic information of the L1 translation equivalents of verbs (Frenck-Mestre & Pynte, 1997; see also Juffs, 2004). For instance, Dussias and Cramer-Scaltz (2008) find that Spanish–English readers are sensitive to verb bias in English, i.e. they learn from experience whether a verb prefers to take a direct object or a sentential complement (see also Lee, Lu, & Garnsey, 2013). Target L2 processing was magnified if verb bias in the L1 and the L2 overlapped and learners had acquired verb preferences in English. These studies illustrate that L2 learners access lexical subcategorization information from the L1 in L2 sentence processing in that L1 verb bias partially guides L2 parsing.

Recent studies using cognates find that non-selective lexical access extends beyond subcategorization information. As reviewed above, cognate words facilitate lexical access since a word form maps to identical conceptual representations in both languages of a bilingual. Focussing on the facilitatory role of cognates for lexical access, Miller (2014) conducted a cross-modal priming experiment on gap filling in the L2 processing of wh-dependencies. Previous self-paced reading and cross-modal priming studies reported that adult L2 readers do not posit (intermediate) gaps in syntactically licensed positions; instead they directly integrate a wh-filler with the verb (Felser & Roberts, 2007; Marinis et al., 2005; yet see Pliatsikis & Marinis, 2013). In cross-modal priming, picture probes that are either related or unrelated to the wh-filler are used in various positions of the sentence to test if and where the wh-filler is reactivated. In a set of experiments, Miller (2014) varied the cognate status of the probes, assuming that cognate probes would allow for faster lexical processing of the probes than noncognates. In object cleft sentences and relative clauses, L1 English intermediate and advanced learners L2 French indeed demonstrated showed targetlike reactivation of nouns in the cognate, yet not the cognate conditions. Miller interprets these
results as reflecting the faster lexical processing of cognates, which, in turn, she argues free resources for computing target syntactic relations.

Facilitatory effects of cognates on lexical access were also demonstrated in a study by Hopp (2017a), building on prior research by Jacob (2009). German-English intermediate to advanced adult learners read sentences as in (5).

(5)  
a. When the doctor Sarah ignored tried to leave the room the nurse came in all of a sudden. (Reduced Relative Clause – Embedded Clause)  
b. The doctor Sarah ignored tried to leave the room when the nurse came in all of a sudden. (Reduced Relative Clause – Main Clause)  
c. When the doctor who Sarah ignored tried to leave the room the nurse came in all of a sudden. (Full Relative Clause – Embedded Clause)  
d. The doctor who Sarah ignored tried to leave the room when the nurse came in all of a sudden. (Full Relative Clause – Main Clause)

If translated word by word, the word order of the preposed temporal clause containing a reduced relative clause (Sarah ignored) in (5a) is temporarily ambiguous with a canonical embedded German clause (Als der Doktor Sarah ignorierte – When the doctor ignored Sarah) due to the SOV order in German embedded clauses. No such syntactic overlap obtains in main clauses (5b), since German has verb-second order in main clauses. In (5c&d), the overt relative pronoun who equally rules out any direct mapping of the English word order onto German syntax. Using self-paced reading, Jacob (2009) found reading delays suggestive of cross-linguistic activation of the L1 syntax among German learners in sentences like (5a). These delays were limited to when readers alternated between reading sentences in English and fillers in German. In eye tracking during reading, Hopp (2017a) extended Jacob’s study by including a cognate manipulation (ignore\textsubscript{COGNATE} versus avoid\textsubscript{NON-COGNATE}) in order to test
whether lexical co-activation and facilitation affect the cross-linguistic activation of L1 syntax in L2 processing.

First fixation durations were significantly shorter for cognates than non-cognates in the L2 group, showing that cognates facilitate lexical processing even in monolingual English sentence contexts. Further, German readers had selective slowdowns on the verb in (5a) in later reading time measures, i.e. first-pass reading times, for sentences with non-cognates, yet not with cognates. These slowdowns obtained for the entire group of German learners in an experiment including German fillers; in a monolingual English-only experiment, only lower-proficiency learners evinced reading time delays reflecting activation of German syntax. In conjunction, the findings demonstrate that cognates modulate cross-linguistic syntactic activation in L2 sentence processing. As in Miller’s (2014) study, the inclusion of cognates leads to more target-like L2 syntactic processing in that L2 readers appeared to have sufficient resources to compute the English target structure and inhibit interference from the L1. Conversely, with non-cognates, lexical processing did apparently not leave enough resources to inhibit the L1 parse. Activation of the L1 is more pronounced at lower and intermediate proficiency levels (see also Hopp, 2017c), and it persist at higher proficiency levels if overall activation of the L1 is heightened in a code-mixing context with L1 filler sentences.

In this selective overview, I summarized a set of recent studies that address the relations between the bilingual lexicon and sentence processing. A growing number of studies finds that lexical processing affects syntactic structure building also in contexts in which lexical information does not encode information relevant for structure-building operations, as is the case with gender for agreement processing or verb bias for complement selection. Rather, properties such as frequency and cross-linguistic form overlap of lexical items incorporated in a sentence contribute to the successful formation of a target parse in L2 processing.
5. The Lexical Bottleneck Hypothesis and models of L2 processing and acquisition

Against the backdrop of these findings, the Lexical Bottleneck Hypothesis claims that some aspects of non-target L2 syntactic processing may be owing to characteristics of bilingual processing in stages that precede and subserve parsing, in particular lexical processing. In consequence, a target-like parse cannot always be effected. Critically, failure by L2 learners to demonstrate target syntactic processing does not mean that the L2 parser or its processes are different from the monolingual parser or that the underlying grammar of the L2 cannot be accessed. Rather, these difficulties can be caused by factors extraneous to syntactic processing as such.

The Lexical Bottleneck Hypothesis, inspired by observations in Dekydtspotter, Schwartz and Sprouse (2006), builds on the functional architecture of the language processing system in which lexical processing precedes syntactic processing and feeds into it. Specifically, the integrated nature of the bilingual mental lexicon with its core characteristics of weaker links and non-selective lexical access can lead to input for syntactic processing that is less robust, more diffuse or delayed. Accordingly, delays or differences in earlier stages of processing that subserve syntactic processing may lead to non-target syntactic processing. In these cases, if difficulties, delays or cross-linguistic influence in lower-level processing are removed or taken into account, adult L2 learners can come to demonstrate target-like syntactic processing in the L2. Conversely, if corresponding difficulties in lower-level processing are experimentally induced in monolingual speakers, L2-like non-target syntactic processing is predicted to ensue.

The functional and temporal staging of lexical and syntactic processing in sentence comprehension is central to the Lexical Bottleneck Hypothesis, and the hypothesis essentially defines their interactions and spells out their consequences for L2 parsing. In consequence, the Lexical Bottleneck Hypothesis by no means aspires to be a model of (L2) sentence
processing, and it certainly does not account for all aspects of non-target syntactic processing, esp. at lower L2 proficiency levels where the interlanguage grammar accessed by the parser may be non-target-like.

In addition, for some phenomena in L2 parsing, effects of lexical processing may be limited or less visible, and lexical processing may be largely irrelevant for others. Even for phenomena, such as gender agreement, for which lexical contributions to non-target syntactic processing obtain, lexical variability does not account for all aspects of non-target processing. Beyond the processing of gender agreement, effects of lexical processing have been found for structural ambiguities and non-local dependencies. For these phenomena, lexical factors often interacted with syntactic complexity and locality. In consequence, effects of lexical processing in parsing appear to be more pronounced in taxing parses, e.g. those requiring substantial reanalysis or the storage and integration of multiple types of information. Conversely, lexical effects may surface even for less complex parses if lexical processing becomes more demanding, e.g. due to the inclusion of low-frequency items.

Besides structural factors, it is necessary to assess how individual factors, e.g. working memory, interact with lexical processing. In Hopp (2014b), lexical processing and reading span did not correlate, yet a systematic survey of individual differences in L2 processing skills awaits investigation. All of these aspects should be tested empirically. As it stands, the Lexical Bottleneck Hypothesis is in need of refinement and limited in scope. However, the core tenets of the hypothesis can be incorporated into existing models of L2 sentence processing and acquisition.

The Lexical Bottleneck hypothesis forms a natural extension to limited-capacity models of non-targetlike L2 sentence processing (e.g. Hopp, 2010; McDonald, 2006; Sorace, 2011; see also Dekydtspotter & Renaud, 2014). According to these models, the computational demands of processing an L2 leave insufficient resources, so that L2 learners cannot integrate all requisite information during on-line comprehension. In consequence, learners often fail to
complete target-like parses and may compute shallow, incomplete or L1-based structures and interpretations. The Lexical Bottleneck highlights how lexical processing depletes resources that curtail syntactic processing, as lexical retrieval impacts the processing of syntactic ambiguities (Hopp, 2014b, 2016b) or agreement relations (e.g. Hopp, 2013; McDonald & Roussel, 2010).

By focussing on the mapping of lexical to syntactic information, the Lexical Bottleneck Hypothesis also becomes relevant in the context of interface approaches to L2 acquisition, e.g. Sorace’s Interface Hypothesis (2011) and Slabakova’s Bottleneck Hypothesis on inflection (Slabakova, 2009). The Lexical Bottleneck Hypothesis adds to these approaches that internal interfaces constitute bottlenecks in the real-time mapping of information across domains and that lexical retrieval difficulties extend beyond the recruitment of inflection.

According to the Shallow Structure Hypothesis (Clahsen & Felser, 2006), adult L2 learners underuse syntactic information versus other information, e.g. lexical-thematic, discourse, plausibility and statistical information, in real-time processing. In its recent version, the Shallow Structures Hypothesis emphasizes that differences between L1 and L2 processing are gradual, rather than qualitative and that they affect the timing of when different information types are used in L1 and L2 processing (Clahsen & Felser, 2017). Part of the reason for the underuse of morphosyntax in L2 learners may be that the time-course of processing mandates L2 learners to assign more relative weight to the processing of lexical information, and, conversely, that signs of the use of syntactic information and syntactic structure building are absent or delayed due to the higher demands of lexical processing. On top of delays in lexical processing having knock-on effects in grammatical processing, differences in word processing may affect later stages of processing. According to the updated version of the Shallow Structure Hypothesis, L2 learners engage in lower degrees of morphological decomposition and rely more on whole-word representations in L2 processing (see also Clahsen & Veríssimo, 2016). Hence, the amount of lexically-based, e.g. inflectional,
information that is accessed during L2 sentence processing may be lower or its retrieval slower, so that it cannot be recruited in syntactic processing. In future research, it will be interesting to investigate whether poor morphological decomposition also entails less detailed syntactic processing.

The relative weighting of different information types forms the central tenet in Cunnings’ (2017) memory-based model of L2 processing. According to Cunnings, L2 speakers suffer from greater cue-based interference and they assign more weights to discourse than to morphosyntactic information in parsing than L1 speakers. Cunnings mentions that one underlying cause for the larger susceptibility of L2 learners to interference may be the quality and extent of lexical processing determining the detail with which lexical items are stored in memory. Less robust lexical representations may be more diffuse and thus give rise to greater interference effects in sentence processing (see also Hopp, 2017b). While relations between lexical knowledge and interference in sentence processing have been found for L1 processing (e.g. van Dyke, Johns, & Kukona, 2014), future studies need to ascertain whether similar relations hold in L2 processing.

6. Conclusion

In this paper, I explored the consequences of the bilingual mental lexicon for L2 sentence processing. Two characteristics of the bilingual mental lexicon, namely weaker lexical links as well as more diffuse lexical activation and representations owing to non-selective lexical access, can give rise to differences in syntactic processing between L2 and L1 speakers. On the one hand, lexical processing may overtax computational capacity, delaying or suspending the application of syntactic structure. On the other hand, weaker lexical encoding and non-selective lexical access reduce the detail of lexical representations, affecting the degree of interference or the relative weighting of lexical versus syntactic information in L2 parsing.
Lexical aspects can thus arbitrate whether, when and how syntactic structure is accessed or applied in L2 sentence processing. In consequence, their consideration merits a central role in research on (L2) sentence comprehension as the field moves towards formulating more comprehensive models of bilingual language processing that comprise phonological, lexical and syntactic processing.

Acknowledgements

I thank the organisers of the 17th Annual Conference of the Japan Second Language Association (J-SLA 2017) conference at Shizuoka University of Art and Culture, Japan, for kindly inviting me to deliver a keynote talk, and I am grateful to the conference participants for comments and stimulating discussion.

References


acquisition of Spanish morphosyntax: The L1/L2 connection (pp. 135-178). Dordrecht: Kluwer.


Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: the effects of L2 knowledge, semantic constraint, and


