

John Benjamins Publishing Company



This is a contribution from *Multilingual Individuals and Multilingual Societies*.

Edited by Kurt Braunmüller and Christoph Gabriel.

© 2012. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Tables of Contents, abstracts and guidelines are available at www.benjamins.com

First exposure learners make use of top-down lexical knowledge when learning words^{*}

Susanne E. Carroll
University of Calgary, Canada

Learning another language requires learning a new lexicon. Current second language acquisition theories make different predictions about the relative importance of L2 experience and L1 knowledge when learning new words. In a study of first exposure learners, clear effects of knowledge of L1 words were found. However, rapid learning after minimal exposure to continuous speech was also found, even when target words contained novel L2 sounds. Results show both the powerful role of L1 lexical knowledge on L2 word learning and the rapid rate at which sound forms are created and mapped to referents. This suggests that a more nuanced approach to discussion of frequency effects and transfer is needed.

Keywords: German, English, second language acquisition, input, word learning, segmentation, transfer

1. Experience and L1 knowledge in L2 word learning

Current second language acquisition theories attribute different degrees of importance to prior experience with the L2. They also attribute different roles to L1-knowledge. Among generativists, Schwartz & Sprouse (1994, 1996) assert that knowledge of the L1 lexicon is the starting point for L2 acquisition, along with L1 parameter-settings. In stark contrast, the Basic Variety model claims there are no L1 effects to be found at the

* I gratefully acknowledge generous funding from the Social Sciences and Humanities Research Council of Canada through its general grants program (grant 410-2006-0323) and its Canada Research Chairs program (grant 950-202408), without which none of this research would have been possible. I thank my research assistants Danica MacDonald, who collected the data, Silke Weber, who recorded the stimuli, and Lindsay Hrats, who has kept the lab running smoothly for many months. I also thank the editors, Christoph Gabriel and Kurt Braunmüller, for the opportunity to present my research in this volume, and two anonymous reviewers whose criticisms and comments have made this a much better paper. Remaining deficiencies are my fault.

initial stage of L2 acquisition (Klein & Perdue 1997). Neither one of these models makes particular claims about input effects. The Basic Variety model assumes it plays a role in accounting for L2 development, while in the Schwartz & Sprouse model, as with all generativist theorizing, the role of input is downplayed because the learner's knowledge is greater than and more abstract than information that could be induced from it (Epstein, Flynn & Martohardjono 1996). The clearest claims about both L1 influence and input effects are to be found in the Competition Model and Construction Grammar usage-based proposals. This is because they are implemented in "bottom-up" feed-forward connectionist architectures. They predict strong across-the-board effects of L1 knowledge on L2 learning because knowledge of language is defined as heavily entrenched associative links between sounds and meanings (Bates & MacWhinney 1987, MacWhinney 2005: 57, Ellis 1998, 2002, 2003).¹ Since connections that have strong weights are easily activated, L2 speech ought to activate L1 lexical representations.² Initial connections between L2 sound forms and L2 lexical representations will nonetheless form on the basis of experience with L2 speech, but they will be weak and will lose out in competition to entrenched connections until such time that the weights in newly-formed connections increase enough to meet some threshold. Changing the weights in connections demands continued L2 listening experience and the belief is that this happens only incrementally (Ellis 1998, 2002, Mellow 2008).

Emergentists are committed to the premise that inputs are physically measurable, objective properties of stimuli (Bates & MacWhinney 1987). Consequently they predict transfer effects during pre-lexical processing (MacWhinney 2005). However, L1 entrenchment effects between sound forms and meanings are not consistent with evidence of rapid word learning at the beginning stage of L2 exposure (Osterhout, McLaughlin, Pitkänen, Frenck-Mestre & Molinaro 2006). An alternative hypothesis is that in learning any word, a learner must construct multiple representations, not just across levels of description (phonology, morpho-syntax, semantics) but *within* each level. Thus, a word might have multiple sound forms, with distinct types being processed as a consequence of different degrees of exposure to input. Learners might initially form L2 representations by segmenting from the speech stream acoustic exemplars. Such representations are rich in acoustic information and are stored in episodic memory (Palmeri, Goldinger & Pisoni 1993, Nygaard, Sommers & Pisoni 1994,

1. MacWhinney (2005: 87) writes: "... the fact that L2 learning is so heavily influenced by transfer from L1 means that it would be impossible to construct a model of L2 learning that did not take into account the structure of the first language."

2. Nothing in connectionist learning models requires that they represent lexical items. MacWhinney (2000) reviews a number of serious problems with models that have dispensed with lexical representations. His discussion brings out clearly that associations between sound forms (either phonetic tokens or phonological types) and meanings must be mediated by a lexical representation. In my discussion of connectionist and usage-based models, I will assume that they include lexical representations, an assumption that also facilitates comparisons between connectionism and algorithm-based learning models.

Goldinger, 1998). Through repeated exposure to specific exemplars in distinct phonetic contexts, more abstract properties such as timing slots (of consonants and vowels), syllables, and phonetic categories, might be computed (MacWhinney & Leinbach 1991). Independently, the acoustic exemplars could be associated to context referents through processing of visual input such as, say, pictures of people and objects. With greater exposure, more abstract representations of meaning might be derived. In this scenario, the learner's representations remain heavily contingent on experience. Consequently, such approaches predict considerable variation among learners based on the frequency of exposure to given words in specific contexts and to the same words in different contexts. They predict important differences among L2ers who have acquired their L2 in tutored contexts, largely through reading written texts and through learner-adjusted speech (Chaudron 1986, Håkansson 1986, 1987), and L2ers whose experience consists exclusively of exposure to speech.

2. Segmenting sound forms, recognizing words and making form-meaning correspondences

When speakers speak, they do not put pauses between their words. Speech is continuous, obscuring the boundaries between words. Nonetheless, when experienced listeners listen, they carve up speech into units of sounds, such as syllables, rhythmic units, and intonational phrases. This process is called *segmentation*. Segmentation roots word recognition in the speech signal. If words are complexes of representations, even learning the sound form of a word may present learning difficulties. Consider a learner who segments a form like ([start]) *start* from the signal, he may or may not recognise *start* when it occurs in ([staɪ]) ([tɪd]) *started* (Dejean de la Bâtie & Bradley 1995, Christophe, Peperkamp, Pallier, Block, & Mehler 2004). Spotting words in larger words is something that proficient users of a language can do (Shillcock 1990, Dumay, Frauenfelder & Content 2002). They make use of phonological, lexical and syntactic information to recognize words (Mattys & Melhorn 2007). In contrast, in learning a first language, infants must segment the speech signal in the absence of such “top-down” knowledge. For example, infants being exposed to English do not know yet that no English word starts with the sequence *pf*. Such phonotactic knowledge will only emerge as the child builds a lexicon. By adulthood, the native English-listener can make use of this knowledge to segment words between the *p* and the *f* (Brent & Cartwright 1996, McQueen 1998). The native German-listener will have had quite different experience since common words like *Pfeffer* ‘pepper’ or *Kopf* ‘head’ reveal that German words can begin and end with *pf*. So segmentation at the initial stage of L1 learning must proceed differently than at the end-state. The infant will rely on recurrent and salient properties of the signal to carve it up into discrete sound units, including rhythm and intonation (Cutler 1994, Jusczyk 1997, Houston 2005).

Learning an L2 presents a situation in-between these two scenarios. On the one hand, on first exposure to another language, the child or adult L2er, like the pre-linguistic infant, has no discrete representations of sound units of the L2. She lacks L2 lexical representations and knowledge of L2 grammar. On the other hand, she brings to the learning task from the L1 much usable knowledge about linguistic units and linguistic structure. She knows, for example, that people can bear proper names and that these are nominal expressions. This kind of knowledge might guide inferential processes during early stages of L2 acquisition and provide information that goes beyond the perceptual salience of proper names in general (Bortfeld, Morgan, Golinkoff & Rathbun 2005). She knows the phonotactic constraints on L1 words. Knowledge of L1 phonotactic constraints has been shown to constrain L2 listening in fluent bilinguals (Avery & Best 1995, Hallé, Segui, Frauenfelder, & Meunier 1998, Tench 2003). When the L1 and the L2 exhibit the same constraints, L2ers presumably draw successfully on L1-knowledge. By hypothesis, if an Anglophone is learning French, where the sequence *pf* is also impossible as a word onset, he will correctly parse the *p* in one syllable and the *f* in the following syllable. What happens in learning German? Presumably some exposure to the novel sequence is required, but how much exposure is needed and under what circumstances is unclear. On the one hand, while L1 phonotactics have been shown to constrain statistical learning on first exposure, learners were still able to segment words beginning and ending with novel combinations (Finn & Hudson Kam 2008). This suggests that experience will result in target-like processing. However, Weber & Cutler (2006) demonstrate that L1 phonotactics continue to constrain L2 segmentation in highly proficient Dutch/English bilinguals. Altenberg (2005), in contrast, found effects of L1 knowledge of Spanish in a production task with intermediate and advanced learners, but not in perception.

Such differences might be due to many factors: to differences in the proficiency levels of the participants, to the language pairs involved (Dutch being in some respects more like English than Spanish); to the fact that the Spanish-speakers were living in an environment where English was the language of broader communication but the Dutch were not; or to differences in the tasks used. More studies that focus on these factors are needed. Another issue that merits attention is the fact that researchers differ in whether they are concerned with the investigation of processing based on extant representations or processing that results in the creation of *novel* representations. There is a fundamental difference between these two concerns that is critical to a coherent discussion of L2 acquisition (cf. Carroll 2001 for discussion). The former asks: How does the learner process speech once she has lexical representations? The latter asks: How does the learner form novel representations in the first place? Studies of learners at different stages of knowledge have traditionally been used to make inferences about the processes that led one knowledge state to turn into a more advanced state. By reducing the time gap between stages of learning, by controlling the input, and by describing it in depth, one can hope to shed real light on the processes that lead to the creation of novel representations. This is the logic behind the research discussed here.

Steady-state L2 users will differ considerably from one another in what they know about the L2 lexicon. While such variation in experience does not preclude studying L2 word learning (cf. Escudero, Hayes-Harb & Mitterer 2008), it requires using tasks that prevent top-down processing and these tasks may lack ecological validity. In other words, while perfectly suited for shedding light on details of speech processing in such knowledgeable listeners (see Weber & Cutler 2004, Cutler & Weber 2007), they are not ideal for drawing conclusions about how learners create novel representations in the first place.

3. Why study first exposure learners?

By choosing participants who have no meaningful exposure to the L2, one can exclude the possibility that some of the learners have encoded representations of the target words through prior experience. By hypothesis, the learners have no L2 lexicon, no knowledge of the L2 sound system and no knowledge of the L2 grammar. Thus, one can exclude top-down processing that draws on L2 representations. This still leaves open the possibility that learners are drawing on L1 representations and L1-attuned processing strategies when analyzing the signal.

What do existing first exposure studies reveal about segmentation and form-meaning mappings?

Numerous studies of segmentation and statistical learning have been carried out that involve presenting adults with a continuous stream of synthesized C(onsonant) V(owel)-syllables from which prosodic cues to word boundaries have been removed. The hypothesis is that listeners will unconsciously compute transitional probabilities (TPs) between syllables and segment words at points of low probability (Saffran, Newport & Aslin 1996). Thus, if a synthesized string includes ... *beragabidumodalidu* ..., where *beraga*, *bidumo* and *dalidu* are all target “words”, exposure will inform the learner that the TPs across *be*, *ra* and *ga* are very high and very low across *ga* and *bi*. When confronted with a word judgement task in which they have to compare sound tokens of target words to sound tokens of “partial words” created by re-aligning syllables across points of higher transitional probability, e.g., ... *be ragabi dumoda lidu* ..., participants prefer the words (Saffran et al. 1996, see Folia, Uddén, De Vries, Forkstam, & Petersson 2010 for a recent review). Results show that segmentation of words under these conditions can occur with as little as two minutes of exposure (Endress & Bonatti 2007). In the absence of typical prosodic cues to word boundaries, this suggests that the computation of TPs plays a role in segmentation. However, a weaker conclusion is motivated. Many of the studies used stimuli consisting exclusively of universal CV-syllables, with the consonants and vowels selected from the L1 repertoire. By

transferring L1 knowledge at the syllable level, the learner ought to be able to prosodically parse such strings. This matters because perception of syllables is presumably a prerequisite for a statistical learning mechanism that counts the distribution of syllables in the input. If the input is, however, unintelligible, that is to say, the learner does not reliably hear syllables, then such a learning mechanism cannot perform. Unfortunately, learners must sometimes learn an L2 where the consonants and vowels are indeed different from those of the L1, and differences in phonetic repertoires are known to create perceptual difficulties for L2ers (Yamada & Tohkura 1992, Lively, Pisoni & Logan 1992, Pisoni & Lively 1995 among many others). Finally, Finn & Hudson Kam (2008: 480) point out that theories of processing and acquisition need to differentiate types of statistical regularities in the input. Transitional probabilities are statistical regularities that are specific to a given speech stream presented during an experiment and might not be stored in memory once a segmented unit has been processed (contrary to the assumptions of Exemplar Theory). In contrast, the statistical computations that are typically discussed in the literature on cue-based learning (and relevant to learning phonotactics) represent computations that must be stored in memory. As noted, Finn & Hudson Kam (2008) showed that constraints on the shape of English prosodic words robustly interfered with statistical learning.

Osterhout et al. (2006) demonstrated that English-speaking tutored learners of French were sensitive to the difference between real words and pseudo-words after only 14 hours of instruction. In an ERP study, they found that pseudo-words reliably elicited an N400 response, even though in a behavioural test, these learners did not reliably distinguish between the two types of words. The authors suggest that this effect may indicate that the learners had rapidly memorized the words, and reported a robust correlation between the N400 word/non-word effect and the frequency of the words in the learners' French textbook. Rapid segmentation and word-learning in a tutored context was also reported by Rast (2008, 2010), who studied native speakers of European French hearing Polish in a communicative language classroom. They were tested on sentence-repetition and translation tests after four and eight hours of instruction. Rast also examined effects of L1 phonological knowledge. Sentences of Polish were presented that contained words that were phonemically similar or not similar to French words.³ Rast differentiated between "frequent" words (presented at

3. They were described as "phonologically transparent" or "not phonologically transparent". To define the phonologically transparent words, an independent group of 15 Francophone adults listened to Polish words and translated them into French. Answers were treated as correct translations if the listener identified a correct Polish root, e.g. *studentem* was translated as *étudiant*, *étude*, *étudiant*. Translations such as *stupéfait*, *stupéfiant*, *soudain*, and "no response" were not accepted as correct. Some of these phonologically transparent words were clearly borrowings. Since Polish is a Slavic language and French is a Romance language, it is unlikely that these words are cognates as defined linguistically (see Carroll 1992 for discussion). In psycholinguistic studies of lexical processing, even nonce forms and words from artificial languages have been described as "cognates" (Van Hell & De Groot 1998, De Groot & Keijzer 2000). In such studies,

least 21 times) and words that occurred less often in the input. After four hours of instruction, learners were able to reliably repeat words from the Polish input but there was no difference between frequent and infrequent input. After eight hours of exposure, participants performed better on the frequent words.

Gullberg, Roberts, Dimroth, Veroude & Indefrey (2010) and Gullberg, Roberts & Dimroth (in press) point out that one should be cautious in generalizing to naturalistic learning contexts from studies of tutored learners. The students in the Osterhout et al. (2006) study were drilled on rules and had discrete orthographic representations of words to aid their word-learning. The French/Polish learners were exposed to auditory input only, but tutored learners are invariably exposed to “teacher talk” with its special adjustments in the use of questions, intonation and recasts to aid comprehension-in-context (Håkansson 1987). The Gullberg et al. studies were designed to study segmentation, frequency and gesture with stimuli that would capture the properties of a particular text-type, namely a televised weather report.⁴ They used a word recognition task, a picture-sound form matching task, and a lexical decision task. “Frequent” words occurred eight times in a text consisting of 120 clauses, “infrequent” words occurred two times, with 292 word types presented in all. Several groups of Dutch native speakers listened to seven or 14 minutes of a Chinese-language videotape. Results showed that participants could rapidly segment words and map them to visual stimuli after seven minutes of exposure. Participants recognized Chinese words better if they were frequent and gesturally highlighted. However, the authors found differences in results that depended on the length of the words; participants were sensitive to distributional properties only with disyllabic words that occurred frequently. Gullberg et al. (in press) attribute this to general difficulties in processing sequences of monosyllables, but it might also indicate an L1-influenced preference for a minimal word of two syllables.

In these studies, participants were asked to watch the videotape without being told what they would see, or why. They also had no idea that they would be tested afterwards. These aspects of the design permit the authors to make strong claims about implicit learning since the participants could not have used encyclopaedic or social knowledge to perform on the various tasks. It does not follow, of course, that inferencing was excluded as the participants listened to the weather report. Given its stereotyped nature, and the fact that the researchers were counting on these properties to aid interpretation, we should conclude that learners probably were using inference to

what matters is that the sound form of one language activates a word of the other language. I shall simplify the discussion by referring to cognates in this way from this point.

4. Thus, their choice of stimuli precluded interaction with the native speaker, a factor that many L2 researchers regard as a *sine qua non* of natural language learning (Gass 1997). Weather reports are also highly conventionalized text types. They draw on an extremely small set of vocabulary items, which are repeated again and again in simple sentence structures. These are just some of the reasons why weather reports were the first text type to be successfully employed in automatic machine translation. See Leplus, Langlais & Lapalme (2004).

bootstrap at least word meaning.⁵ In short, this design too had its limitations, a fact that should encourage us to use a number of different designs to shed light on word learning at the initial state.

4. Our studies

4.1 Methodology and stimuli

We have carried out a number of studies on L2 word learning using a paradigm that presents controlled auditory input. Learners were trained in two phases: participants were told that they would hear sequences of speech in German while looking at visual stimuli in the form of line drawings of people. They were told to learn the names of the individuals they saw. They then listened to single declarative sentences presented one after the other. As each sentence was heard, a different picture was presented on a computer screen. These were followed by 20 single questions and the same line drawings. The questions presented the listeners with a forced-choice between two options; examples are given below.⁶ We were interested in studying if adults could accurately pick out the target name and map each name to the correct visual referents. By hypothesis, the task presupposes that the participants can segment the target name from continuous speech.

What is interesting about our paradigm, in comparison to those described above is that we are able to measure responses to inputs in terms of number of exposures, based on the learner's own behavioural responses. Crucially, we measure increased exposure, not in terms of an arbitrary period of time (7 minutes, 4 hours, 14 hours), but in terms of participants' performance on the task. In addition, we are the first to add a developmental aspect to the investigation of first exposure learners by testing participants approximately two weeks after initial training. We are thus able to answer the question: Do participants retain in long term memory representations of the words

5. This should be obvious on the basis of a "thought experiment". Imagine that instead of talking about the sunrise and sunset while pointing to a picture of a sun on the weather map the speaker had recorded the following: "Don't you just love the colour yellow. Doesn't it just make you feel good when you see the sun come up." The whole point of using a weather report was to ensure the inference that the talk be about the objects on the map, rather than the decontextualized talk that typifies much of our language use.

6. All stimuli were standardized, completely randomized, and the pace was experimenter controlled. The experiments were programmed in E-prime, a platform for computer-based experiments. Subjects were tested individually in the laboratory while seated comfortably at a computer terminal. They heard auditory stimuli through AKG k 171 studio-quality headphones. They had 2500 ms. to respond. Response keys were clearly marked on the keyboard with subjects pressing the F1 key if they thought the correct response was the first name in the NP conjunct, or the F12 key if they thought the correct response was the second name in the NP conjunct.

they were exposed to, especially after a hiatus with no input? Or, are the initial representations ephemeral and quickly disappear?

The design of the studies is presented in Table 1. I will present in detail the results from Study 2 – the cognate/non-cognate comparison. Cognate data from Studies 1 and 3 will be presented afterwards since they amount to replications of the cognate Exp. C with different groups of participants.

Table 2 presents an overview of the procedures used.

Although our task was less ecologically valid in some respects than that of Rast (2008), where interactions between the learners and a native speaker occurred, or Gullberg et al. (2010, in press), where the learners heard 7 minutes of continuous text, we did attempt to create stimuli that would be as natural as possible.⁷ Proper names were chosen because they are very often among the first words that an L2er learns. The fact that we have found no published data dealing with proper names in L2 acquisition suggests that it is widely assumed that they will be easy to learn.⁸ This

Table 1. The studies (*Order of the experiments was counter-balanced across subjects in each study, e.g. for half of the participants, Exp A < Exp B; for the other half Exp B < Exp A)

Study 1* L1 vs. L2 comparison	Exp. A: English first names in English sentences	Exp. B: German first names in German sentences (all names were cognate items)
Study 2* Cognate vs. non-cognate comparison	Exp. C: Same stimuli as Exp. B	Exp. D: German first names in German sentences (all names were non-cognate items)
Study 3* Length comparison	Exp. E: Same stimuli as Exp. B	Exp. F: German cognate and non-cognate first names + German last names in German sentences

7. Gullberg et al. (2010, in press) report on the lexical properties of their stimuli but not on the phonetic or phonological properties. It is inconceivable the 7 minutes of continuous text equals 7 minutes of continuous speech. Pauses undoubtedly occurred in the stimuli, at the very least, at the ends of sentences. If so, our stimuli are perhaps more like those of the Gullberg et al. studies, with the difference that we have conducted phonetic analyses of our stimuli and can provide descriptions of the breaks in phonation, changes in fundamental frequency, syllable duration, and amplitude.

8. Bortfeld et al. (2005) show that familiar names are highly salient for infants learning their L1. It seems reasonable to assume that familiar names will also be salient to L2ers. Eavesdropping on conversations in an unfamiliar language should convince the reader that familiar names of politicians, actors, or shopping malls will “pop out” from the rest of the L2 speech. Monaghan and Christiansen (2010) suggest that proper names, which may occur in the linguistic input without other linguistic context, once initially represented, may provide an easy route to segmenting stretches of speech lying in-between the name and the right edge of the utterance.

Table 2. Procedures used**Session 1**

1. Ethical consent procedures were completed (subjects were provided with a general description of the purpose of the study).
2. Background questionnaire (questions related to language of the home, number of languages known, time spent outside of Canada, etc.)
3. Digital span task
4. Instructions in English (presented auditorily and in writing) along with practice items to familiarize subjects with the response method. Participants were told that they would hear sentences in German while looking at pictures of people, e.g. *Hier sehen Sie Reinhold*. 'Here you see Reinhold.' They were to learn the names of the people depicted. Participants were also told that they would hear questions asking them about the people depicted and were to press the F1 key if the correct response was the first word of two presented in a question, and to press the F12 key if the correct response was the second of two words presented in a question. E.g. *Sehen Sie hier Reinhold_{F1} oder Reinhart_{F12}*? 'Do you see here Reinhold or Reinhart?' (Practice performed at computer console was self-paced.)
5. Training session (1 to a maximum of 10)
Participants heard 20 statements and saw 20 different pictures. (Timing was experimenter-controlled). They then heard 20 questions and saw the same pictures again. (All stimuli were randomized; participants had 2500 msec. to respond.)
At the end of each statement/question cycle, the accuracy score (Measure 1) of the participant appeared on the screen. If the participant scored 20/20, they proceeded to Test phase. If not, the training session was repeated until the participant got all items correct (Measure 2), or the experiment was terminated.
6. Test
Participants saw a different picture of the same individuals.
After each decision on the receptive task, participants pronounced their choice. Production data were recorded.

Session 2 (c. 2 weeks later)

7. Re-test1
Participants saw the same pictures as in the Training sessions. The position of the names was reversed from that appearing in the Training and Test questions. Feedback was given if the participants made an error.
After each decision on the receptive task, participants pronounced their choice. Production data were recorded.
8. Re-test2
The pictures and questions of Re-test1 were used.
After each decision on the receptive task, participants pronounced their choice. Production data were recorded.

assumption might turn out to be correct for various reasons, but it merits further investigation.⁹

The proper names chosen are typical German first names.¹⁰ Cognate names were chosen because it was hypothesized that they might activate L1 names. Non-cognate names were chosen on the assumption that they would require the learners to create novel sound forms both phonetically and phonologically. In each experiment, 20 target names were embedded in one of four declarative sentence frames. Each frame is a natural way to introduce new information and includes deictic elements like *hier* ‘here’, *da* ‘there’, or *das* ‘that’ that referred to the pictures. Examples of both cognate items (C) and non-cognate (NC) items are shown in (1):

- (1) a. Hier ist Agnes. (C)
 [hi:ɪstʔagnəs]
 Here is Agnes.
- b. Da steht Claudia. (C)
 [daʔtɛtklaudia]
 There stands Claudia.
- c. Hier sehen Sie Lutz. (NC)
 [hi:zɛ:ɛnzilɔts]
 Here see you Lutz.
- d. Das ist Annegret. (NC)
 [dasistʔanəgrɛt]
 That is Annegret.

A single native speaker of Standard German produced all recordings. Sentences were between three and nine syllables long (mean = 5.4 in cognate sentences; mean = 5.1 in non-cognate sentences, $t(19) = 0.6305$, $p = 0.53$, n.s.). Each declarative sentence presented participants with an acoustically unique stimulus. After hearing all 20 declarative sentences, participants heard the corresponding questions to each presentational

9. A reviewer suggested that proper names are not “real” words because they are not “meaningful”. Proper names are, to the contrary, good words since they have prosodic structure, and belong to a grammatical class. In fact, proper names can be shown to occur in most of the syntactic contexts where common nouns occur (Jonasson 1994). Moreover, proper names have a meaning which is what permits them to refer and to combine with predicates in sentences like *Not every Susanne has blue eyes*. It is true, that expressions like *Susanne* in their naming function are neither predicates nor descriptions. As Katz (2001) puts it, proper names do not have “senses” which is why they do not license inferences from the particular properties of the INDIVIDUAL referred to other INDIVIDUALS who bear the same name. Not every person bearing the name *Susanne* has blue-eyes. Katz analyses their meaning as “the unique *x* who bears the name [...]” where “...” is either a sound form or an orthographic representation.

10. Typical does not mean “fashionable”. We did not use borrowings from English like *Kevin*, *Patrick* or *Peggy*.

frame with the target item embedded in a conjoined noun phrase. Target words occurred either as the first noun in the conjunct (in utterance-medial position) or as the second noun (in utterance-final position). These questions were also all acoustically unique. At no point did any of the subjects see written versions of the auditory inputs. See (2).

- (2) a. Ist hier Dietmar oder Detlef?
Is here Dietmar or Detlef
- b. Steht da Helga oder Heidrun?
Stands there Helga or Heidrun
- c. Sehen Sie hier Lutz oder Ludo?
See you here Lutz or Ludo
- d. Ist das Annika oder Annegret?
Is that Annika or Annegret

The input on Measure 1 (accuracy on Training Trial 1) consisted of 40 distinct instances of speech. These 40 instances were repeated in subsequent training trials. A participant who learnt all 20 items on a single training trial would have been exposed to a particular declarative sentence exactly once. He would have been exposed to the corresponding questions once on Training Trial 1 and twice on Test. A participant who learnt all 20 items on 10 training trials would have been exposed to a particular declarative sentence 10 times and to the corresponding questions 11 times. This allows for a clear test of the prediction that segmentation and sound-form mappings occur incrementally and require repeated exposure to instances of speech in order to create appropriate associative connections. If processing at the phonetic level depends on L1-entrenched connections, we predict that segmentation of the words will be incremental and slow. Consequently, few or no learners should correctly respond to all 20 items on Training Trial 1 and we expect performance to improve significantly as more training trials are performed. We also predict relatively poor performance in Re-test 1. This is because two weeks later participants were presented with 20 new distinct acoustic exemplars for which no input-output mappings have been established. This is because we changed the positions of the target nouns.¹¹ Alternatively, if participants are analyzing L2 speech using abstract prosodic constituents such as syllables and feet, and if they are analyzing the input in terms of the phonetic categories of the L1, performance on Training Trial 1 might be very good, and frequency of exposures to the same instances of speech might not matter to forming sound-meaning mappings. On this scenario, performance on Re-test 1 should also be good since the learners will be able to segment a word even when it appears in a new

11. In other words, if the target was the first noun in the training and test questions, it became the second noun in the retest phase; if the target was the second noun during training and test, it became the first noun in the retest phase.

position in the input. Exposure might matter, however, to list learning. In other words, individual items might be readily segmented and mapped to referents depicted in the line drawings but learning all 20 items might require repeated exposure to fix the representations in long term memory. Finally, if learners use only the categories of the L1 to process the input, they should show no evidence of having represented unique L2 sounds in their production data. Instead, we should find L1 lexically-triggered pronunciations of the cognates and L1-accented pronunciations of the non-cognates.

4.2 Participants

Participants in all studies were English-speaking students from the University of Calgary who were paid a small fee for their participation. To qualify for the first exposure experiments, participants had to indicate on our background questionnaire that they had no prior knowledge of German, and had not been systematically exposed to German at home or elsewhere. We began with 33 participants in Study 2, 10 of who were not able to learn all 20 items in 10 training trials.¹² 23 participants performed the Test; 21 returned for the re-test phase of the study.

4.3 Results

Table 3 presents mean accuracy scores on the cognate and non-cognate items on Training Trial 1, Test, Re-test 1 and Re-test 2, as well as the mean number of trials to reach criterion.

Table 3. Study 2, cognate vs. non-cognate items

Word type	Training Trial 1	Test	Re-test 1	Re-test 2	# of Trials to criterion
Cognates	82.7%	95%	85.9%	93%	3.04
Non-cognates	70.6%	95%	79%	90%	5.56
Mean	81.1%	95%	82.5%	91.5%	4.3

12. All of these participants were tested first on the non-cognate items. The mean accuracy score of those who did not finish the non-cognate experiment was .66 on Trial 1 versus .719 for those who did, $t(31) = -1.30$, $p = 0.20$, n.s. We submitted all of our participants to a digital span task prior to conducting the main experiments. A comparison of the digital span scores of the subjects who did not finish versus those who did showed no difference between the groups (4.13 vs. 4.1). I have no explanation for why 10 participants had trouble learning the non-cognate items, especially in the light of a later study which showed that participants could readily learn much longer names consisting of a first + last name, e.g. *Rainer Weisskopf* or *Gabrielle Blauhemd*.

One-sample means comparison tests revealed that performance on each measure was well above chance for each word type.¹³ As Table 3 shows, participants had higher mean accuracy scores on the cognate words on Training Trial 1. A two-sample t-test on the difference of these means was significant, $t(55) = -4.09$, $p = 0.0001$. All stimuli had been presented in random order. Analyses of the responses according to the order of presentation of the stimuli in Training Trial 1 showed no correlation between accuracy and position of the item during the training trial. In other words, participants were as accurate in their responses on the first 5 items (before they would have been exposed repeatedly to the different frames) as they were on the last 5 items. Table 3 also shows that participants reached criterion faster on cognate than on non-cognate items. A paired t-test on the difference in the number of trials to criterion showed that this difference was significant, $t(22) = -5.61$, $p = 0.000$.

Table 3 shows that once all items were acquired, the type of word involved no longer made any difference. When re-tested two weeks later, however, participants selected the correct name when prompted by the question significantly better when it was a cognate name, $t(40) = 2.74$, $p = 0.009$. Once they had received feedback on the correct response, type of name again no longer had an effect on selection of the correct answer.

Table 4 shows the mean accuracy scores on each Training Trial as well as the number of individuals who scored 100% on each trial. As this table shows, participants reached criterion sooner on the cognate items; three participants learnt all items on Trial 1, almost half had learnt all items in two trials, and more than two-thirds had learnt all cognate items in three trials; by the end of Trial 5, only one participant had not learnt all items. In contrast, acquisition of the list of non-cognate items increased incrementally.

Table 4. Performance on the training trials

	Cognates Mean	# of 100% scores	Non-cognates Mean	# of 100% scores
Trial 1	82.7%	3	70.6%	0
Trial 2	92.3%	7	79.8%	1
Trial 3	93.5%	6	85.3%	4
Trial 4	95%	4	86.7%	1
Trial 5	97.5%	3	88.8%	5
Trial 6	90%	0 (N = 1)	91.1%	4
Trial 7	95%	0 (N = 1)	93.3%	5
Trial 8	100%	1	91.1%	1
Trial 9	–	–	92.5%	2
Trial 10	–	–	92.5%	–(N = 10)

13. For cognates – Training Trial 1: $t(23) = 36.89$, $p = 0.000$; Test: $t(23) = 78.40$, $p = 0.000$; Re-test 1: $t(20) = 50.59$, $p = 0.000$; Re-test 2: $t(20) = 62.55$, $p = 0.000$. For non-cognates – Training Trial 1: $t(32) = 30.97$, $p = 0.000$; Test: $t(22) = 74.49$, $p = 0.000$; Re-test 1: $t(20) = 35.26$, $p = 0.000$; Re-test 2: $t(20) = 49.27$, $p = 0.000$.

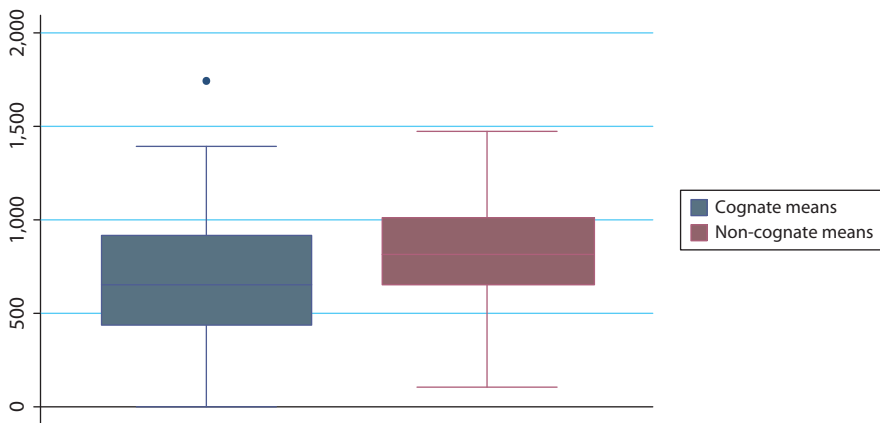
Table 5. Cognate comparisons (Studies 1–3)

Study	Training Trial 1	Test	Re-test 1	Re-test 2	# of Trials to criterion
1	85.6% (N = 25)	96% (N = 23)	87.9% (N = 22)	93.6% (N = 22)	3.21
2	82.7% (N = 33)	95% (N = 23)	85.9% (N = 21)	93% (N = 21)	3.04
3	78.3% (N = 26)	95.7% (N = 26)	87.7% (N = 22)	96.3% (N = 22)	3.88
Mean	82.2%	95.6%	87.1%	94.3%	3.37

Table 5 present the results of the cognate stimuli from all three studies, and shows that the results are robust across three different groups of Anglophones.

To sum up the accuracy results, recall that our participants had no representations of either the German cognates or the non-cognates. Nevertheless, participants were segmenting both kinds of target words after only two exposures to the words in distinct acoustic inputs. In addition, we found robust effects of prior lexical knowledge in that our participants had higher scores on Training Trial 1 and Retest 1 on cognate items. Frequency of the input seemed to affect the ability to correctly associate a long list of 20 items to the appropriate pictures but not the ability to correctly map specific words to their pictures.

Response latencies were calculated from the onset of the stimulus to the offset.¹⁴ Data are presented for correct responses only. Figure 1 shows that responses were

**Figure 1.** Response latencies from Study 2

14. Calculating latencies from the beginning of the stimulus to the end resulted in a negative value whenever the respondent pushed the key immediately after NP1, which they frequently did when NP1 was the correct response.

much faster on cognate items (681.7 ms, $N = 90$) than on non-cognate items (818.33, $N = 98$). This difference was significant, $t(186) = -2.94$, $p = 0.003$. The latency data provide strong evidence that the cognate stimuli were activating L1 lexical entries.¹⁵

Finally, although space limitations preclude detailed discussion of the production data (see Carroll 2010b), the analysis of transcriptions from Study 2 revealed that

Table 6. Qualitative and quantitative analysis of production data (stressed syllables are capitalized)

Name#	Word Type	Ssyll	Error type	Wksyll 1	Error type	Wksyll 2	Error type	Wksyll 3
1 KAI	NC	100.0%		–				
2 HEIke	NC	100.0%		94.7%				
3 HEIdrun	NC	100.0%		42.1%				
4 DIETmar	NC	100.0%		77.3%	Vowel + R			
5 aNIta	C	100.0%		50.0%	Vowel	70.80%		
6 reGIIna	NC	100.0%		50.0%	R	100%		
7 caroLIIna	C	95.8%	Vowel	0.0%	Vowel	0%	R	100%
8 JoHAnna	C	95.8%		91.7%	Onset	100%		
9 JOsef	C	95.5%		36.4%	Vowel, Coda			
10 Eberhardt	NC	95.5%		31.8%	Vowel + R			
11 geSIIne	NC	95.2%		95.2%		100%		
12 SENta	NC	90.0%		100.0%				
13 BENno	NC	86.4%	Vowel	100.0%				
14 FEMke	NC	81.8%	Vowel	95.5%				
15 Eva	C	73.9%	Vowel	52.2%	Vowel, Onset			
16 CLAUdia	C	66.7%	Vowel	100.0%		100%		
17 GEorg	C	66.7%	Vowel	58.3%	Vowel, Coda, R			
18 SÖNke	NC	58.8%	Vowel	100.0%				
19 HARald	C	58.3%		12.5%	Vowel, Coda			

15. It must be emphasized that the fact that our subjects were responding with [ˈbeno] *Benno* when they heard this non-cognate word does not show that *Benno* failed to activate English *Ben*, *Bennie*, *Benji* etc. Given the results here and in Rast (2008) it is highly likely that all L2 words that are similar enough to L1 words will activate those words. However, *Benno* is presumably not “similar enough” to *Bennie* for the speaker to say [ˈbeni]. What phonetic and phonological properties make a word similar or not similar enough is a crucial question in the study of cognates.

participants were very good at repeating the target words. However, they were much more target-like on the non-cognate items than on the cognate items. See Table 6.

Lessened accuracy in pronouncing the German cognate words was due to the fact that they were often pronounced like the corresponding English words. Still some participants attempted to produce some aspect of the cognate-items input. So [jozɛf] *Josef* was not pronounced with an English affricate [dʒ], and [ɛfa] *Eva* was not pronounced as [iv] *Eve* but with [f] and the vowel [e]. Of course, the non-cognate items were produced with English-influenced gestures too, however, there was considerable variation observed and some participants were remarkably target-like on the (non-L1) mid front rounded vowels of *Sönke* and *Jörg* [œ] which would be inexplicable if the subjects had not encoded correct acoustic detail for these items. Thus, pronunciation of these words showed that learners were drawing on different types of sound representations – a representation that preserves information from the signal, as well as representations that encode English-only sounds.

5. Discussion and conclusions

Learning another language entails learning a new lexicon. Words can be seen as a triplex of autonomous representations (Jackendoff 2002a, 2002b). In particular, we may surmise that the learner must segment sound forms from the speech signal, learn the morpho-syntactic properties of a word (its grammatical class, person, number, or gender features, its transitivity, etc.), and its semantic properties (its sense and its referent in a particular context). There is good reason to assume that acquiring all of these properties will demand considerable experience with the L2. Listening to an L2 is known to be more difficult for L2 learners and fluent bilinguals than for native speakers. Even advanced L2ers misparse words when listening to continuous speech (Voss 1977); they are disadvantaged when listening to the L2 in noise (Florentine 1985, Takata & Nabelek 1990, Hardison 1996, inter alia); they have difficulty adjusting to less familiar accents (Major, Fitzmaurice, Bunta & Balasubramanian 2002). Emergentist theories of second language acquisition predict such findings because they predict strong L1 transfer effects even during pre-lexical processing. Such effects might arise because L1 representations are impeding analysis of the input, or the simultaneous activation of both L1 and L2 representations might lead to difficulties in integrating information from multiple sources. The fact that our learners were able to encode for both receptive and production purposes the non-cognate words suggests that entrenchment at pre-lexical levels of processing is not an absolute barrier to segmentation and word learning. These results are consistent with the findings of other studies of first exposure learners (Osterhout et al. 2006, Finn & Hudson Kam 2008, Rast 2008, 2010, Gullberg et al. 2010, in press). We have succeeded in narrowing down even further the amount of experience the L2 requires: not much.

Our study provides clear evidence of L1-based lexical effects when first exposure learners are listening to a second language. These results are consistent with strong transfer theories, including not only connectionist models but also the Full Transfer/ Full Access theory of Schwartz & Sprouse (1994, 1996). Nevertheless, the fact that some first exposure learners produced cognate words with input-based sounds (such as [efa] and not [iv] 'Eve') shows that L1-lexical activation is not deterministically connected to the learners' pronunciation. As well, the fact that at least some learners attempted to pronounce novel sounds in the non-cognate words in a target-like way suggests that these participants have encoded a rich phonetic representation that preserves information from the signal, as exemplar-based theories predict. Because our learners had no prior exposure to German, we can reject as an account of such learner behaviours the hypothesis that some of the learners had already acquired these German names.

Our subjects performed well above chance on our task even on the first training trial, after only two exposures to the proper names. The mean score on the non-cognate items was 70.6%, substantially above the mean scores of 55% reported by Gullberg et al. (in press) after one exposure to the weather report and 60% after two exposures, but comparable to means of 68% reported for two-syllable items that were frequent. This is encouraging and suggests that our stimuli and task are capturing similar learning effects. However, because our paradigm allows us to directly measure the effects of cumulative exposure, we can say that hearing more of the same input had little effect on the processing of individual utterances in the input and that learners can rapidly memorize segmented stimuli (phonetic tokens) on far less input than had previously been supposed (Osterhout et al. 2006). This still leaves open the question of how much exposure is needed to learn a more abstract representation of the sound form of a word which would be needed to recognize the same word when pronounced by different talkers or by speakers with different accents.

Our data show not only that our participants can associate the sound forms of the words with a meaning (a referent), they can do so when the visual cues to the referent change and after a two-week period with no input at all. This is consistent with the findings of Gullberg et al. (2010, in press) and runs counter to an idea expressed in Osterhout et al. (2006: 224) that first exposure learners first learn a sound form and only subsequently, given more experience with L2 input, derive a meaning. Such findings show the necessity of basing claims about input effects on studies of input that are causally related to specific behavioural or other outcomes.

Our results strongly suggest that discussions of frequency in L2 acquisition should take account of the nature of the learning problem. It has been argued here that word-learning entails learning multiple representations that must ultimately be linked in memory as diverse instances of "the same word". Our results show that repeated exposure to the same input is not necessary to memorize a sound token. This is an interesting result in the light of studies of statistical learning in which the same stream of word sequences is presented again and again in a continuous loop (Saffran et al. 1996). It

also adds greater precision to studies that have similar findings but measured exposure in terms of minutes or hours of time. If our interpretation of the data is correct, it follows that repetition of the same tokens will not be useful to a learning mechanism that is sensitive to *changes* in syllable sequences in the input. This is presumably part of the story as to why beginner L2ers segment recurrent strings of words as if they were a single phonetic or phonological unit (Wong Fillmore 1976, Wray 2002, Carroll 2010a). It might well be involved in explaining why our task is so easy in comparison to, e.g., listening to a weather report. One could make the case, as one reviewer did, that it is “obvious” that our task is easy because the stimuli make use of recurrent words. But, as noted, this fact is probably not relevant. Rather, it may be the case that the task is easy because the input can be readily parsed into feet. If our first exposure learners indeed represented the signal in terms of abstract L1-specific metrical categories such as feet and prosodic words (Cutler 1994), then they could have parsed the stimuli into sequences of one or two prosodic words before the utterance-medial targets and two prosodic words after, e.g. [d²sɪst] ... [odɐ] ... *das ist ... oder ...* where the “...” stand in for the proper names. See Carroll (2011) for just such an analysis of the input. However, it must be emphasized that if first exposure learners imposed feet and prosodic words onto the first utterances they heard, this is a significant finding and not something to be presupposed.

Cutler & Shanley (2010: 1844) suggest that listening to L2 speech is “inordinately hard” and may require targeted training. Our study suggests that this is not necessarily so. When the L1 and the L2 are similar in terms of their repertoires of phonetic categories, their abstract phonological categories, and how those categories are instantiated in lexical items (the cognates), there are clear processing advantages for the first exposure learner. However, even when the L1 and the L2 differ in their consonants and vowels, their syllable structures, their stress patterns, and their use of tones (as is the case in the Dutch/Mandarin Chinese study of Gullberg et al. 2010, in press, or the Polish/French studies of Rast 2008, 2010), learners can still create discrete representations of sound tokens on the basis of minimal exposure.

To conclude, not all aspects of L2 learning require repeated exposure to the same inputs and not all aspects of L2 learning are slow and incremental. This is good news indeed because it shows that commonly held beliefs of many L2 learners (and possibly even SLA researchers) that listening to an L2 is hard may underestimate their true initial capacities. Controlled input studies with first exposure learners offer a feasible approach to the investigation of initial processing capacities, and are already shedding interesting light on what the learner can do with the input she receives.

References

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

- Avery, R. A. & C. T. Best. 1995. Phonological and phonotactic influences on perception of two non-native vowel contrasts. *Journal of the Acoustical Society of America* 97: 33–62.
- Bates, E. & B. MacWhinney. 1987. Competition, variation, and language learning. In *Mechanisms of Language Acquisition*, ed. B. MacWhinney, 157–193. Hillsdale NJ: Lawrence Erlbaum Associates.
- Bortfeld, H., J. Morgan, R. M. Golinkoff & K. Rathbun. 2005. Mommy and me: Familiar names help launch babies in speech-stream segmentation. *Psychological Science* 16: 298–304.
- Brent, M. R. & T. A. Cartwright. 1996. Distributional regularity and lexical access. *Journal of Psycholinguistic Research* 26: 363–375.
- Carroll, S. E. 1992. On cognates. *Second Language Research* 82: 93–119.
- Carroll, S. E. 2001. *Input and Evidence: The Raw Material of Second Language Acquisition* (Language Acquisition and Language Disorders 25). Amsterdam: John Benjamins.
- Carroll, S. E. 2010a. Explaining how learners extract ‘formulae’ from L2 input. *Language, Interaction and Acquisition/Langage, Interaction et Acquisition* 1: 229–250.
- Carroll, S. E. 2010b. Exploring the locus of transfer in the L2 pronunciation of novel words. Paper presented at the Annual Meeting of the Canadian Linguistic Association, Concordia University, Montréal, Canada, 31 May 2010.
- Carroll, S. E. 2011. What properties of input influence word learning on first exposure. Paper presented at Eurosla 21, Stockholm University, 8–10 September 2011.
- Chaudron, C. 1986. The role of simplified input in classroom language. In *Learning, Teaching and Communication in the Foreign Language Classroom*, ed. G. Kaspar, 99–110, Aarhus: Aarhus University Press.
- Christophe, A., S. Peperkamp, C. Pallier, E. Block & J. Mehler. 2004. Phonological phrase boundaries constrain lexical access: Adult data. *Journal of Memory and Language* 51: 523–547.
- Cutler, A. 1994. Segmentation problems, rhythmic solutions. *Lingua* 92: 81–104.
- Cutler, A. & J. Shanley. 2010. Validation of a training method for L2 continuous-speech segmentation. In *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010)*, 26–30 September 2010, 1844–1847, Makuhari, Chiba, Japan.
- Cutler, A. & A. Weber. 2007. Listening experience and phonetic-to-lexical mapping in L2. In *Proceedings of the 16th International Congress of Phonetic Science (ICPhS 2007)*, eds. J. Trouvain & W. J. Barry, 43–48. Saarbrücken: Universität des Saarlandes.
- De Groot, A. M. B. & R. Keijzer. 2000. What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning* 50: 1–56.
- Dejean de la Bâtie, B. & D. C. Bradley. 1995. Resolving word boundaries in spoken French: Native and non-native strategies. *Applied Psycholinguistics* 16: 59–81.
- Dumay, N., U. H. Frauenfelder & A. Content. 2002. The role of the syllable in lexical segmentation in French: Word-spotting data. *Brain and Language* 81: 144–161.
- Ellis, N. C. 1998. Emergentism, connectionism and language learning. *Language Learning* 48: 631–664.
- Ellis, N. C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24: 143–188.
- Ellis, N. C. 2003. Constructions, chunking, and connectionism: The emergence of second language structure. In *The Handbook of Second Language Acquisition*, eds. C. J. Doughty & M. H. Long, 63–103. Oxford: Blackwell.

- Endress, A. & L. L. Bonatti. 2007. Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition* 105: 247–299.
- Epstein, S. D., S. Flynn & G. Martohardjono. 1996. Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences* 19: 677–758.
- Escudero, P., R. Hayes-Harb & H. Mitterer. 2008. Novel second-language words and asymmetric lexical access. *Journal of Phonetics* 36: 345–360.
- Finn, A. & C. L. Hudson Kam. 2008. The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition* 108: 477–499.
- Florentine, M. 1985. Speech perception in noise by fluent, non-native listeners. *Proceedings of Inter-Noise 1985*, 1021–1024.
- Folia, V., J. Uddén, M. de Vries, C. Forkstam & K. M. Petersson. 2010. Artificial language learning in adults and children. *Language Learning* 60: 188–220.
- Gass, S. M. 1997. *Input, Interaction and the Second Language Learner*. Mahwah NJ: Lawrence Erlbaum Associates.
- Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251–279.
- Gullberg, M., L. Roberts & C. Dimroth. In press. What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics*.
- Gullberg, M., L. Roberts, C. Dimroth, K. Veroude & P. Indefrey. 2010. Adult language learning after minimal exposure to an unknown natural language. *Language Learning* 60: 5–24.
- Håkansson, G. 1986. Quantitative aspects of teacher talk. In *Learning, Teaching and Communication in the Foreign Language Classroom*, ed. G. Kaspar, 83–98. Aarhus: Aarhus University Press.
- Håkansson, G. 1987. *Teacher Talk: How Teachers Modify their Speech when Addressing Learners of Swedish as a Second Language*. Lund: Lund University Press.
- Hallé, P. A., J. A. Segui, U. Frauenfelder & C. Meunier. 1998. The processing of illegal consonant clusters: A case of perceptual assimilation? *Experimental Psychology: Human Perception and Performance* 24: 1–17.
- Hardison, D. M. 1996. Bi-modal speech perception by native and non-native speakers of English: Factors influencing the McGurk effect. *Language Learning* 46: 3–73.
- Houston, D. M. 2005. Speech perception in infants. In *The Handbook of Speech Perception*, eds. D. B. Pisoni & R. E. Remez, 417–448. Malden MA: Blackwell.
- Jackendoff, R. 2002a. *Foundations of Language, Brain, Meaning, Grammar, Evolution*. Oxford: OUP.
- Jackendoff, R. 2002b. What's in the lexicon? In *Storage and Computation in the Language Faculty*, eds. S. Nooteboom, F. Weerman & F. Wijnen, 23–58. Dordrecht: Kluwer.
- Jonasson, K. 1994. *Le nom propre: Constructions et interprétations*. Louvain-la-Neuve: Éditions Duculot.
- Jusczyk, P. W. 1997. *The Discovery of Spoken Language*. Cambridge MA: The MIT Press.
- Katz, J. J. 2001. The end of Millianism: Multiple bearers, improper names and compositional meaning. *Journal of Philosophy* 98: 137–166.
- Klein, W. & C. Perdue. 1997. The basic variety (or: Couldn't natural languages be much simpler?) *Second Language Research* 13: 301–347.
- Lepus, T., P. Langlais & G. Lapalme. 2004. Weather report translation using a translation memory. In *Machine Translation: From Real Users to Research. Proceedings of the 6th Conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC*,

- USA, *September/October 2004*, eds. R. E. Frederking & K. B. Taylor, 154–163. Berlin: Springer.
- Lively, S. E., D. B. Pisoni & J. S. Logan. 1992. Some effects of training Japanese listeners to identify English /r/ and /l/. In *Speech Perception, Production and Linguistic Structure*, eds. Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka, 175–196. Amsterdam: IOS Press.
- MacWhinney, B. 2000. Lexicalist connectionism. In *Models of Language Acquisition: Inductive and Deductive Approaches*, eds. P. Broeder & J. Murre, 9–32. Oxford: OUP.
- MacWhinney, B. 2005. New directions in the Competition Model. In *Beyond Nature-nurture: Essays in Honor of Elizabeth Bates*, eds. M. Tomasello & D. I. Slobin, 81–110. Mahwah NJ: Lawrence Erlbaum Associates.
- MacWhinney, B. & J. Leinbach. 1991. Implementations are not conceptualisations: Revising the verb learning model. *Cognition* 40: 121–157.
- Major, R. C., S. F. Fitzmaurice, F. Bunta & C. Balasubramanian. 2002. The effects of non-native accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly* 36: 173–190.
- Mattys, S. L. & J. F. Melhorn 2007. Sentential, lexical, and acoustic effects on the perception of word boundaries. *Journal of the Acoustical Society of America* 122: 554–567.
- McQueen, J. M. 1998. Segmentation of continuous speech using phonotactics. *Journal of Memory and Language* 39: 21–46.
- Mellow, D. 2008. Emergentist approaches to teaching and measurement: It's about time. Paper presented at the Annual meeting of the Canadian Association of Applied Linguistics, University of British Columbia, Vancouver, BC, 4 June 2008.
- Monaghan, P. & M. H. Christiansen. 2010. Words in puddles of sound: Modelling psycholinguistic effects in speech segmentation. *Journal of Child Language* 37: 545–564.
- Nygaard, L. C., M. S. Sommers & D. B. Pisoni. 1994. Speech perception as a talker-contingent process. *Psychological Science* 5: 42–46.
- Oosterhout, L., J. McLaughlin, I. Pitkänen, C. Frenck-Mestre & N. Molinaro. 2006. Novice learners, longitudinal designs, and event-related potentials: A means for exploring the neurocognition of second-language processing. *Language Learning* 56: 199–230.
- Palmeri, T. J., S. D. Goldinger & D. B. Pisoni. 1993. Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19: 1–20.
- Pisoni, D. B. & S. E. Lively. 1995. Variability and invariance in speech perception. A new look at some old problems in perceptual learning. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research*, ed. W. Strange, 433–459. Timonium: York Press.
- Rast, R. 2008. *Foreign Language Input: Initial Processing*. Clevedon: Multilingual Matters.
- Rast, R. 2010. The role of linguistic input in the first hours of adult language learning. *Language Learning* 60: 64–84.
- Saffran, J. R., E. L. Newport & R. N. Aslin. 1996. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35: 606–621.
- Schwartz, B. D. & R. Sprouse. 1994. Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops (Language Acquisition and Language Disorders 8)*, eds. T. Hoekstra & B. D. Schwartz, 317–368. Amsterdam: John Benjamins.

- Schwartz, B. D. & R. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12: 40–72.
- Shillcock, R. 1990. Speech segmentation and the generation of lexical hypotheses. In *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*, ed. G. T. M. Altmann, 24–49. Cambridge MA: The MIT Press.
- Takata, Y. & A. K. Nabelek. 1990. English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America* 88: 663–666.
- Tench, P. 2003. Non-native speakers' misperception of English vowels and consonants: Evidence from Korean adults in the UK. *International Review of Applied Linguistics in Language Teaching (IRAL)* 41: 145–173.
- Van Hell, J. & A. M. B. de Groot. 1998. Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition* 1: 193–211.
- Voss, B. 1977. *Hesitation Phenomena as Sources of Perceptual Errors for Non-native Speakers* (Applied and interdisciplinary papers, Series B 25). Trier: L.A.U.T.
- Weber, A. & A. Cutler. 2004. Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language* 50: 1–25.
- Weber, A. & A. Cutler. 2006. First-language phonotactics in second language listening. *Journal of the Acoustical Society of America* 119: 597–607.
- Wong Fillmore, L. 1976. *The Second Time Around: Cognitive and Social Strategies in Second Language Acquisition*. PhD dissertation, Stanford University.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: CUP.
- Yamada, R. A. & Y. Tohkura. 1992. Perception of American English /r/ and /l/ by native speakers of Japanese. In *Speech Perception, Production and Linguistic Structure*, eds. Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka, 155–174. Amsterdam: IOS Press.