The Aspect Hypothesis and the Acquisition of L2 Past Morphology in the Last 20 Years

Kathleen Bardovi-Harlig, Indiana University
ORCID ID 0000-0003-4058-9318

Llorenç Comajoan-Colomé, University of Vic – Central University of Catalonia
ORCID ID 0000-0001-9815-674

Abstract

Twenty years ago, a state-of-the-art review in *SSLA* marked the coming of age of the study of temporality in second language acquisition. This was followed by three monographs on tense and aspect the following year. This paper presents a state-of-the-scholarship review of the last 20 years of research addressing the aspect hypothesis (AH) (Andersen, 1991, 2002; Andersen & Shirai, 1994, 1996), the most tested hypothesis in L2 temporality research. The first section of the paper gives an overview of the AH and examines its central tenets, and then explores the results of empirical studies that test the hypothesis. The second section considers studies that have investigated four crucial variables in the acquisition of temporality and the testing of the AH. The third section discusses theoretically motivated areas of future research within the framework of the hypothesis.

Twenty years ago, *SSLA* published a review of tense-aspect research in second language acquisition, "From morpheme studies to temporal semantics" (Bardovi-Harlig, 1999). The following year saw the publication of three monographs on L2 temporality (Bardovi-Harlig, 2000; Li & Shirai, 2000; Salaberry, 2000), which Slabakova (2002) described as both a maturing of the field and an opportunity for additional research. The number of studies of the acquisition of temporality and tense and aspect has grown dramatically in the last two decades. In this review, we present a critical synthesis of recent research addressing the aspect hypothesis (AH) (Andersen, 1986, 1990, 1991, 2002; Andersen & Shirai, 1994, 1996), the single most influential hypothesis in second language acquisition (SLA) research regarding tense and aspect.

Research in the last 20 years can be characterized as having tested the universal against the particular. The universal is represented by the aspect hypothesis, and the particular by proficiency, first language, input, and task. The overarching question has been whether the aspect hypothesis holds, or whether the acquisitional sequences it predicts can be disrupted by individual variables.

As L2 research on the AH developed, researchers provided clear operationalization for future research, breaking the AH down into easily testable hypotheses (Shirai, 1991; Andersen & Shirai, 1996), simplifying tests for identifying aspectual categories (Shirai, 2013), describing calculations (Bardovi-Harlig, 2000, 2002), and discussing research methodology pertaining to the study of tense-aspect (Salaberry & Comajoan, 2013; Salaberry & Shirai, 2002). The increased accessibility of the AH has made it a popular topic of investigation, and in this review, we assess the accumulated empirical support for the AH, discuss the impact that four recurrent variables have on tests of the AH and our ability to interpret them, and propose theoretically-motivated questions for future research.

The Aspect Hypothesis

The Evolution of the Aspect Hypothesis

The AH builds on three main constructs: tense, grammatical aspect, and lexical aspect. Tense establishes the location of an event in time (Comrie, 1985). Grammatical aspect allows for "ways of viewing the internal temporal constituency of a situation" (Comrie, 1976, p. 3).[1] Grammatical aspect is sometimes called *viewpoint aspect* (Smith, 1983), because the choice between progressive and simple, or perfective and imperfective, for example, often reflects the speaker's view of the action (Table 1).

Table 1. Grammatical Aspect in the Past in Three Language Families

Grammatical aspect in the past (morphology)	English: *sing*	Japanese: uta-*u*	Romance: Spanish: cant-ar French: chant-er Italian: cant-are
Perfective	Ana *sang*.	Ana-ga utat-*ta*.	Sp: Ana cant-*ó*. Fr: Ana *a* chant-*é*. It: Ana *a* cant-*ato*.
Imperfective	Ana *was* sing-*ing*.	Ana-ga utatt-*te* i-*ta*.	Sp: Ana cant-*aba*. Fr: Ana chant-*eait*. It: Ana cant-*ava*.
Morphological Progressive	Ana *was* sing-*ing*.	Ana-ga utatt-*te* i-*ta*.	Sp: Ana *estaba* cant-*ando*. Fr: none It: Ana *stava* cant-*ando*.

For instance, in English, a contrast in grammatical aspect is found between simple past and progressive (cf. "Ana sang" and "Ana was singing"). A similar distinction is found in Japanese. In Romance languages, the contrast between perfective and what is called general imperfective is found only in the past: the preterite (in Spanish), *passé composé* (in French), and *passato prossimo* (in Italian) are the perfective forms; and the imperfect (in Spanish), the *imparfait* (in French), and the *imperfetto* (in Italian) are general imperfective forms. Whereas perfective morphology, in general, presents a situation from the outside and has a terminative interpretation, imperfective morphology views the situation from the inside and has an unbounded interpretation, with different functions (habitual, progressive, or continuous; Deo, 2012; De Swart, 2012). Spanish and Italian (among others) also have a morphological progressive, which like the progressive in English, can occur in different tenses. French, however, lacks a morphological progressive, but progressivity is a concept expressed by the imperfect. Grammatical or viewpoint aspect additionally allows a speaker to portray events for rhetorical purposes (such as the main action of a narration, or as auxiliary information to it).

In contrast, lexical aspect (also known as *inherent aspect*, *Aktionsart*, and *actionality*, Binnick, 1991) is nongrammatical and refers to semantic differences in verbs and their arguments (Dowty, 1979) or *verb constellations* (Smith, 1997), such as whether a predicate has inherent duration like "talk" and "sleep," is punctual like "recognize" and "notice," or has elements of both duration and culmination like "build a house" and "paint a picture."[2] This may be most clearly illustrated with predicates such as "swim" and "swim a mile." "Swim" (intransitive) is an activity, whereas "swim a mile" is an accomplishment. "Swim" has no inherent endpoint, whereas "swim a mile" does, namely the end of the mile. Dowty (1979) used different tests to distinguish lexical aspectual categories. One category is the adverbial test with

"for/in" + time: "for" + time distinguishes activities, "He swam for 30 minutes," from accomplishments, #"He swam a mile for 30 minutes"); "in" + time is acceptable with accomplishments "He swam a mile in 30 minutes" but not with activities #"He swam in 30 minutes," where "30 minutes" is the duration of swimming. Similarly, the completion test also works to distinguish activities from accomplishments: If someone is swimming and she is interrupted after 15 minutes, did she swim? The answer is "yes." But if someone who swims a 30-minute mile is interrupted after 15 minutes, did she swim a mile? The answer is "no." Thus, we see that lexical aspect belongs to the predicate and not the verb (Binnick, 1991; Dowty, 1979; Klein, 2009; Shirai, 2013).

In its simplest form, the AH for SLA predicts that in the initial stages of the acquisition of tense-aspect morphology by adults, the acquisition of past morphology will be influenced by lexical aspectual categories.[3] Namely, verbal morphology will be attracted to and will occur with predicates with similar semantics. Perfective past will occur with telic predicates (predicates with inherent endpoints); in contrast, imperfective will occur with unbounded predicates, and progressive will occur with ongoing activities.

The AH in SLA research (Andersen & Shirai, 1994; Bardovi-Harlig, 1994) underwent a series of refinements in fairly rapid succession in response to empirical evidence. The earliest version of the AH appeared as the defective tense hypothesis (Andersen, 1986, 1991, following Weist et al., 1984) (Figure 1):

In beginning stages of language acquisition only *inherent aspectual* distinctions are

encoded by verbal morphology, not tense or grammatical aspect (Andersen, 1991, p.

307; emphasis in the original).

Figure 1 about here

_____

Opposing tense and grammatical aspect appeared to be too strong because emerging morphology was shown to mark past tense as well as favoring congruent lexical aspect (Bardovi-Harlig, 1992; Robison, 1995; cf. Bloom et al. 1980). The next formulation of the aspect hypothesis (Andersen & Shirai 1994, 1996) —using the name by which we now know the aspect hypothesis—maintained the importance of the initial influence of aspect, but did not explicitly oppose aspect to tense:

> First and second language learners will initially be influenced by the inherent semantic aspect of verbs or predicates in the acquisition of tense and aspect markers associated with or/affixed to these verbs (Andersen & Shirai, 1994, p. 133).

In 1996, Andersen and Shirai translated Andersen's (1991) figure into text and explicitly linked Vendler's (1957/1967) account of lexical aspect—arguably the dominant theoretical account of aspect at the time (see also Dowty, 1979)—to the AH by stating the predicted developmental stages in terms of the four Vendler categories: States, activities, accomplishments, and achievements (see Figure 1; see Shirai, 1991, pp. 9-10 for first language acquisition).[4]

1. Learners first use (perfective) past marking on achievements and accomplishments, eventually extending use to activities and statives.

2. Imperfective past appears later than perfective past, and imperfect past marking begins with statives, extending next to activities, then to accomplishments, and finally to achievements.

6

3. In languages that have progressive aspect, progressive marking begins with activities, then extends to accomplishments and achievements.

4. Progressive markings are not incorrectly overextended to statives (Andersen & Shirai, 1996, p. 533).

Andersen and Shirai (1994, 1996) claimed that the Vendler categories were linguistic and cognitive universals in the sense that they could be found in a large number of languages and that they mirrored cognitive and perceptual human distinctions, thus giving the AH the status of a semantic and cognitive universal that accounted for its effect both in L1 and L2 acquisition.

Linking the AH to Vendler categories allowed the hypothesis to be presented as four easily testable hypotheses, instead of one larger, but less specific hypothesis. It made the hypothesis more testable by spelling out the relationship between specific aspectual classes and specific grammatical aspects.

In his last article on the AH, entitled "The Dimensions of 'Pastness'," Andersen (2002) again positioned the AH unambiguously in the expression of the past— "The Aspect Hypothesis in its simplest form makes a clear prediction for first and subsequent uses of past forms" (p. 79)—and formulated three subhypotheses.[5]

1. Learners first use past marking (e.g., English) or perfective marking (Chinese, Spanish, etc.) on achievement and accomplishment verbs, eventually extending its use to activity and then to stative verbs.

2. In languages that encode the perfective-imperfective distinction, a morphologically encoded imperfective past as in Romance languages appears later than perfective past,

and imperfect past marking begins with stative and activity verbs, then extends to accomplishment or achievement verbs.

3. In languages that have progressive aspect, progressive marking begins with activity verbs, then extends to accomplishment or achievement verbs (p. 79).

Hypotheses 1 and 2, as stated in Andersen (2002), make less of a distinction between the Vendler categories than Andersen and Shirai (1996) did. In Hypothesis 1 achievements are grouped with accomplishments (telics), and in Hypothesis 2 statives are grouped with activities (atelics). Thus, even cases in which accomplishments exhibit greater use of perfective than achievements may be considered evidence in favor of the AH.

The fourth hypothesis (posited in Andersen & Shirai, 1996) was eliminated because L2 acquisition showed occasional use of progressive with states: "progressive markers are sometimes overextended to stative verbs in some of the L2 studies" (Andersen & Shirai, 1996, p. 544).[6] Andersen (2002) reduced the eight stages of development from Figure 1 (Andersen, 1991) to two lines showing the development of morphology (Figure 2). The figures differed in three ways: First, Figure 2 no longer depicts 8 stages of the acquisition of perfective-imperfective morphology, but rather two paths of development (one for perfective and one for imperfective). Second, the depiction of the development of morphology in Figure 2 was simplified in that imperfective emerges later than posited in Figure 1, rather than immediately after perfective emerges in achievements at stage 2. Figure 2 suggests that imperfective emerges only after the perfective spread is well underway, as Andersen (2002, p. 83) explained: "Learners do not use Past Imperfect at all for a long time, even while they are progressing in their use of Preterit on more and more verbs." Third, the use of present as an imperfective marker preceding the use of

8

the past imperfect (Andersen, 1991, 2002; Bardovi-Harlig, 2005; Kaplan, 1987) is no longer depicted in Figure 2. Neither figure illustrates the path of the progressive posited in Hypothesis 3.

_____

Figure 2 about here

_____

A Review of 20 years of SLA Research on the Aspect Hypothesis

Identifying Empirical Studies for the Review

This review of research on the AH spans 1999 (the publication of Bardovi-Harlig's 1999 review) to 2019, 20 years later. It follows a narrative analysis rather than a statistical analysis and considers the perspectives as complementary (cf. Ellis, 2018). For the compilation of SLA research in the AH, we adopted some of the methods in systematic reviews (e.g., EPPI-Centre, 2006). First, we conducted a search of SLA and applied linguistics journals from 1999 to 2019 using the key words *tense, aspect, acquisition of morphology,* and *aspect hypothesis*, employing both journal and publisher search engines. We also manually searched the published proceedings of tense-aspect conferences. Second, once the initial list of articles had been compiled, we decided to concentrate exclusively on empirical research on the AH. We then manually searched the articles to verify that they reported an empirical study in which the AH was directly addressed in the research questions or the goal statements of the articles, either by invoking the name of the hypothesis or by a paraphrase of the hypothesis or one of the subhypotheses. Finally, we also considered the references of articles that we were working with, and added any sources that had not been identified in the original computer searches. Since we were focusing on the

development of L2 temporal morphology by adults, studies on L1 acquisition and child L2 acquisition were excluded. Doctoral dissertations were also excluded unless they were subsequently published as articles or books. In this way, we identified 36 studies published between 1999 and the third quarter of 2019 in 13 journals, 7 edited volumes, and two monographs, establishing a set of studies that test the AH which, if not exhaustive, is at least representative. (See Table 1 in the online supplementary materials for a tabular summary of the studies under review.)

Review of Empirical Studies about the AH in SLA: 1999-2019

In this section, we provide a brief overview of studies that address the AH, first considering studies that claim to support it, and next considering studies that claim to show counterexamples.

Bardovi-Harlig (1999) described the breadth of research in tense-aspect from multiple perspectives, including meaning-oriented and form-oriented perspectives. When discussing the AH she concluded that it was "widely supported" (p. 362) and suggested what would constitute counterevidence to the AH:

> The clearest counterexample to any formulation of the aspect hypothesis would be an interlanguage system that exhibits equal distribution of verbal morphology in all categories—that is, states, activities, accomplishments, and achievements. That would involve the emergence of the preterit, for example, appearing equally in all categories; the same for present, imperfect, and progressive as they emerge. Counterevidence would not necessarily involve contrast between morphemes at the earliest stages...

The matter of what constitutes "equal distribution of verbal morphology in all categories" is dependent on the aspectual categories that are implemented in any given study, because AH studies initially used binary aspectual distinctions and only later incorporated the four Vendler categories. For instance, authors of early studies considered the evidence from binary distinctions to support the AH (e.g., Kaplan, 1987, tested events vs. non-events which are telic vs. atelic; and Robison, 1990, tested both the punctual-durative contrast and the stative-dynamic contrast). Although the four-way Vendlerian categorization has been widely adopted, three- and five-category analyses have also been proposed (e.g., Kenny, 1963; Smith, 1997; see also fn. 4). In this paper, we adopt the view that the AH is about the influence (or lack thereof) of the lexico-aspectual properties of verbal predicates in the emergence of past morphology (including both perfective and imperfective morphology), and we consider the binary telic-atelic distinction to be evidence for the AH, because the binary distinctions also reflect lexical aspectual categories.

The first and second hypotheses of the AH are the most well researched and are often studied together. Evidence for Hypotheses 1 is found in the distribution of a greater percentage of perfective forms (e.g., past in English, preterite in Spanish, *passé composé* in French, past "-ta" forms in Japanese) emerging in telic than in atelic lexical categories (i.e., emergence in achievements and accomplishments). Evidence for Hypothesis 2 is found in a pattern of emergence and development of imperfective morphology from states and activities to accomplishments and achievements (i.e., earlier emergence with atelic predicates of imperfect forms in Romance languages). Of the 29 studies that have investigated how perfective and imperfective forms emerge (Hypothesis 1 or Hypothesis 2, or both), 23 reported results that they found to be consistent with the AH: 17 found that emergence of morphology was closely linked to lexico-aspectual categories (target languages include Catalan, French, Italian, Portuguese,

Spanish, and for Hypothesis 1 only, English), and an additional 6 reported that the evidence partially supported the AH (target languages include Chinese, English, French, and Korean). (See Table 1 in the online supplementary materials). Among the earliest studies to offer empirical evidence for the development of perfective and imperfective both in written and oral data in the early stages of acquisition were Camps (2005) and Comajoan (2006). Camps (2005) provided evidence for the AH in the emergence of both perfective (preterite) and imperfective (imperfect) morphology by studying written language production of true beginner students of L2 Spanish (first-year learners at the university level without any instruction in Spanish in high school). The results of the study showed a clear lexico-aspectual pattern in the emergence of preterite and imperfect: 83.9% of achievements marked with preterite, 67.7% of accomplishments, 63.4% of activities, and 39.1% of states; compared to 47.8% of states marked with imperfect, 29.0% of activities, 22.6% of accomplishments, and 10.7% of achievements in Composition 2, p. 170). Comajoan (2006) investigated beginner learners of Catalan whose oral production showed that there was a clear association of perfective morphology with telics and imperfective with atelics.

In the case of Hypothesis 1, results that are inconsistent with the AH raise two main issues: a) statives may be marked for past tense earlier than other categories (Ayoun & Salaberry, 2008; Camps, 2005; Zhao & Shirai, 2018; Domínguez et al., 2013; also reported by earlier studies such as Bardovi-Harlig & Reynolds, 1995, and Collins, 2002, among others), and b) perfective forms may be used in all lexico-aspectual categories (e.g., Labeau, 2005; Salaberry 2000; 2008). Regarding Hypothesis 2, results that are inconsistent with the AH raise issues of spread (or lack thereof) across categories (imperfective morphology is associated with states and not with

activities; Domínguez et al., 2013; Izquierdo & Kihlstedt, 2019; imperfective spreads across all categories; Salaberry, 2005[7]).

That states attract past earlier than predicted has been reported by multiple studies. Without competition from another grammatical past form in English, the simple past appears early on statives (in contrast to Romance languages, in which the contrast between perfective and imperfective is realized only in the past). English shows clear evidence of spread of perfective past from achievements to accomplishments to activities, but the use of past with statives seems to be accelerated compared to the prediction made by Hypothesis 1; that is, the appropriate use of past with statives can surpass activities (Bardovi-Harlig & Reynolds, 1995; Collins, 2002; Ayoun & Salaberry, 2008). Several factors seem to be in play regarding the inflection of stative predicates, including limited stative vocabularies (Bardovi-Harlig, 2005), the fact that "be" (the most dominant stative) almost always appears tensed (not in base form), and a lack of competition from other past forms found in the other lexical aspectual categories. In Romance languages, with perfective-imperfective past morphology pairs (e.g., Spanish "ser,*"* imperfective, "era*"*; perfective, "fue*"*; French "être," *imparfait,* "était*"*; *passé compose,* "a été*"*), statives seem to follow a different path; they are inflected earlier both in imperfective (as predicted by the AH) and perfective morphology. For instance, Domínguez et al. (2013) showed that, in the oral data of the least proficient learners, preterite and imperfect morphology use in L2 Spanish was similar for states (around 20% average use). Salaberry (2008) discussed similar results for low proficiency learners of Spanish.

Some studies have argued against Hypotheses 1 and 2 of the AH stating either that telicity was not the main variable that accounted for use of morphology, but rather other lexico-semantic distinctions (e.g., dynamicity, Domínguez et al. 2013; González & Quintana, 2018) or that

13

learners initially use a default past tense marker across all lexical categories (i.e., the default past tense hypothesis, Salaberry, 2000, 2008).

Domínguez et al. (2013) argued that their data showed that dynamaticity and not telicity influenced the use of tense-aspect morphology, but in fact their data show mixed results regarding the AH. Whereas the group with the lowest proficiency in L2 Spanish (Year 10 of lower secondary school in England, with 200 hours of instruction) had a similar rate of preterite use in states and achievements (contrary to AH), one of the categories (namely, accomplishments) received the highest rate of preterite use. The data from Year 13 (750 hours of instruction) showed that there was no difference in the use of preterite in the dynamic categories (achievements, accomplishment, and activities), and thus it was considered contrary to the AH. However, in the undergraduate advanced learner group (900 hours of instruction), the data showed a clear effect of lexical class, since rates of preterite in the four categories were significantly different.

In a second set of data, Domínguez et al. (2013) examined the use of imperfect in a personal narrative, an impersonal narrative with a habitual frame ("every day was the same…") followed by an action frame ("and then one day…"), and a controlled narrative, which elicited non-prototypical combinations of lexical aspect and morphology by supplying pictures with predicates and adverbs. The results showed that use of imperfect tended to follow the expected aspectual pattern in the less controlled narratives, but it did not in the controlled narrative, where the beginner and intermediate learners showed use of imperfect only in states or in states and accomplishments (the use of preterite was not reported).

Even though they acknowledged the influence of task, discourse structure, and learner proficiency in their results, Domínguez et al. (2013, p. 12) interpreted their data as

counterevidence for the AH: "In the light of all these results, we have enough evidence to argue that a pattern of emergence and development of past tense forms across different lexical classes consistent with the L[lexical]AH is not supported by our corpus of data." Such a strong claim, however, needs to be questioned for two main reasons. First, their data did show, as acknowledged by the authors themselves, clear effects of lexical aspect in the use of morphology (e.g., in advanced learners, with use of preterite significantly different in the four categories). However, the clear aspectual pattern was not evident for all learner groups and tasks. Second, even though there was not a clear distinction in the use of morphology according to the four categories, there was in most cases at least a distinction between telic and atelic predicates (Domínguez et al., 2013, Tables 4 and 5, p. 8). In sum, the data from Domínguez et al., (2013) provided inconclusive evidence regarding the AH and in some respect support it if we consider the binary telic-atelic distinction.

González and Quintana (2018) investigated the acquisition of L2 Spanish by L1 English and L1 Dutch learners, and found that their data could be better explained by a durative vs. terminative contrast (Verkuyl, 1993), which they equate to the telic-atelic contrast (see footnote 4 in González & Quintana, 2018), than by the four-category Vendlerian classification. In this respect, the authors claimed that "the L[lexical]AH needs to be revised… There is no influence of the Vendler classification in the L2 production of our students" (p. 622). However, this seems to be an overstatement. A reanalysis of González and Quintana's (2018) across-category analysis actually distinguishes the different Vendlerian categories in the use of imperfect and telic-atelic in the perfective (see "Methodological Issues").

A second type of counterevidence to the AH is provided by Salaberry (2008, p. 13), who posited the default past tense hypothesis:

During the first stages of L2 development, learners will attempt to mark tense distinctions relying on a default morphological marker (most typically the perfective form), whereas inherent semantic aspect and discursive factors are expected to have an increasing effect as learners' proficiency improves throughout time…

The evidence for the default past tense hypothesis comes mostly from the fact that beginner learners of L2 Spanish (L1 English) inflected states with perfective morphology inappropriately in the early stages of acquisition, which resulted in similar rates of use of perfective morphology across aspectual categories (contrary to the AH; see Comajoan, 2006 for a discussion of results). Although the default past tense hypothesis was posited to account for the early acquisition of past morphology, Salaberry (2008) reviewed a number of studies that showed that advanced learners may follow it as well. Even though the default past tense hypothesis is presented in opposition to the AH (e.g., Domínguez et al., 2013; Gonazález & Quintana, 2018), Salaberry (2008, p. 187) acknowledged that in the case of beginner learners, "the rationale that underlies the proposal of a default marker of past tense is also based on basic notions underlying the LAH and the DH [discourse hypothesis]." In sum, in our view, the default past tense hypothesis does not contradict the AH, but rather attempts to explain the acquisition of specific pairings of languages (L2 with perfective-imperfective contrast, L1 without, e.g., Spanish-English), in specific learning environments (mostly, classrooms, where perfective forms are often presented first and in an isolated way) and for specific proficiency levels (mostly, beginners).

Hypothesis 3, concerning the spread of past progressive, has been less studied than the other hypotheses. Evidence comes from languages in which the progressive is expressed via specific morphology (e.g., progressive in English, Japanese, or Korean), as opposed to languages where progressive meaning can be expressed by general imperfectives (e.g., in Romance languages).

16

Overall, evidence suggests that Hypothesis 3 holds since progressive past morphology is more frequent in activities than in the other lexico-aspectual categories. Bardovi-Harlig (2012) investigated the progressive in a longitudinal study of L2 English and found that the use of progressive morphology mostly occurred with activities with a continuous reading. The progressive next spreads to iterative or repeated readings within the same categories. Although Bardovi-Harlig observed that the development of the progressive had interim stages that precede the spread to accomplishments and achievements, development was consistent with the AH. In contrast, Muñoz and Gilabert (2011) claimed to present counterevidence to Hypothesis 3, reporting that progressive forms were more frequently used with accomplishments than with activities (contrary to the AH). However, the difference was not significant and there was a clear task effect. In addition, they did not use the proportional analysis used by other studies. Instead, they used an across-category analysis, which is sensitive to the number of predicates produced in each aspectual class (see section "Methodological Issues"). Muñoz and Gilabert also reported that their picture task favored the production of accomplishments, thus further calling into question their interpretation that greater use of progressive with accomplishments is meaningful as it was task induced, and uncontrolled for in the across-category analysis.

Studies have also investigated the development of progressive morphology in L2 Japanese (Sugaya & Shirai, 2007; Ryu, Horie, & Shirai, 2015) and L2 Korean (Kim, 2012; Lee & Kim, 2007). The results of these studies draw a general picture of progressive morphology associated with activities, influenced by a number of additional factors (L1, language-specific patterns, U-shaped behavior, instruction, input frequency, and learner characteristics). These results led Ryu, Hori and Shirai (2015, p. 816) to conclude: "The real issue is perhaps not whether the Aspect Hypothesis holds in all situations, but under what conditions ... Considering this, it would be

wise to treat the descriptive generalization of the Aspect Hypothesis as a 'universal tendency,' not as an 'absolute universal'." Before we discuss the additional factors investigated within the AH (section 2), we first focus on methodological issues that determine how results are interpreted.

Methodological Issues in SLA Research on the Aspect Hypothesis

This section discusses three methodological issues that are central to how data are analyzed and interpreted as evidence or counterevidence for the AH: a) the unit of analysis, b) data reporting, and c) within- versus across-category data analyses.

The unit of analysis for lexical aspectual category is the predicate, not the verb (Dowty, 1979), and this has important consequences for how data are quantified and interpreted. Consider accomplishments, which are often compositional as in "build a house," "read a book," and "write a paper." The same verb takes different arguments, and changing the predicate changes the lexical aspect as well: "Read a book" is an accomplishment, but both "read" and "read books" are activities. Thus, studies that analyze verbs rather than predicates show a mismatch of linguistic units of analysis with the concept of lexical aspect, and thus the AH. The studies that most conspicuously base their analyses on verbs rather than predicates are corpus studies (e.g., Daidone, 2018; Tracy-Ventura & Cuesta-Medina, 2018; Wulff et al., 2009).

The unit of analysis is also related to the different lexico-aspectual categories that are taken as basic for the analysis of data. As mentioned earlier, what tests are used and how categories are established varies (e.g., telic vs. atelic; stative vs. dynamic, and so on), and it will necessarily affect study results. A further issue is whether such analyses can be made confidently in interlanguage, illustrated by the question, "how do we know what learners know?" (Lakshmanan

& Selinker, 2001; Lardiere, 2003; Shirai, 2007). In this regard, Shirai (2013) identified a number of methodological problems in assigning a lexical category to predicates produced by learners, and Rastelli and Vernice (2013) and Rastelli (2019) have argued that L2 learners in the beginning stages do not assign lexical aspect to verbal predicates and that when they do so they may do it structurally (using frequent verbs and adjuncts) but not lexically.

The way in which studies report data is also important. Some studies fail to report the data in its entirety, which jeopardizes our ability to understand the process of language acquisition. This happens in at least two ways: failure to report all the forms that learners produce or failure to report on all the learners who participated. Reports that include the full range of tense-aspect forms produced by learners, thus providing an interlanguage analysis, reveal the range of expression of a single temporal concept (e.g., Collins, 2002, 2004; Comajoan, 2006; McManus, 2013; Sugaya & Shirai, 2007; Uno, 2014; Vraciu, 2013). Learners produce a variety of tense-aspect forms even in past time contexts, which include (for English) past, past progressive, present, base, and other forms, including present perfect (Bardovi-Harlig, 2000; Bardovi-Harlig & Reynolds, 1995). In contrast, other studies report only three forms (Diaubalick & Guijarro-Fuentes, 2017; Howard, 2004) or two (the preterite and imperfect, Domínguez, et al., 2013; González & Quintana, 2018; Izquierdo & Collins, 2008; Labeau, 2005).[8] These reports fail to show how developing verbal morphology fits into the larger system. Muñoz and Gilabert (2011) report only progressive production and exclude learners who did not produce progressives, thus inflating what learners appear to be able to do.

A similar issue arises from eliminating responses to items. Diaubalick and Guijarro-Fuentes (2017) reported that 112 learners completed the production test, but they did not report how the 112 learners were distributed across proficiency levels. Nevertheless, the number of items should

be constant for the entire group. However, almost 50 responses (15%) per category are missing.[9] If the discrepancies arise from excluding items that learners have left blank, giving an analysis of only those items that learners could answer, this inflates the description of learner abilities.

Other studies report only *correct* scores in SOC (supplied in obligatory context) analysis (Ayoun & Salaberry, 2008; Labeau, 2005; Zhao & Shirai, 2018). Sugaya and Shirai (2007) reported the results of a picture description task by providing SOC scores (pp. 21-23), and they also provided a supplementary interlanguage analysis, reporting "distribution of alternative forms." The second analysis is important because the AH is an interlanguage hypothesis, and as such rests on what learners produce, not only on what they produce well. Moreover, the suppliance of all forms provides a better test of the AH as Collins's (2004) analysis of French learners' use of present perfect shows. The present perfect in English is not simple past, but it is a perfective past for the L1 French learners. Thus, by using present perfect with achievements and accomplishments, learners follow Hypothesis 1. This is an important example of interlanguage analysis, and a response to the creeping "correct use" orientation in AH research.

Distinguishing within-category and across-category analyses of tense-aspect distribution is crucial in assessing whether learner production is consistent with the AH, because if one aspectual class has more tokens, it will have a higher percentage of use of morphology. The most common analysis in research on the AH is a within-category analysis, which calculates the distribution of morphology within each category regardless of the number of tokens of each aspectual category (Bardovi-Harlig, 1999, 2000, 2002; Bardovi-Harlig & Reynolds, 1995; Collins, 2002, 2004; Comajoan, 2006; Camps, 2005; Dominguez et al., 2013; Salaberry, 2002, inter alia). Comparisons of the within and across analyses exist in the published literature (e.g., Bardovi-Harlig, 2000, 2002; Labeau, 2005). For instance, Bardovi-Harlig (2000, 2002)

recalculated the across-category analyses of Rohde (1996) and Salaberry (1999, 2000), respectively, to demonstrate that across category analyses that were claimed to provide counterevidence to the AH, instead support the AH when presented as within-category analyses (see also Sugaya & Shirai, footnote 9, p. 31). In contrast, across-category analyses are presented by Chan et al. (2012), Fessi (2016), González and Quintana (2018), and Muñoz and Gilabert (2011).

As an illustration, Table 2 provides the original across-category analysis for Dutch-speaking learners of Spanish from González and Quintana (2018). The arrow that runs left to right indicates that the total is calculated across categories. The learners produced 43 achievements and an equal number of activities (44) and more than twice as many accomplishments (96). The occurrence of preterite on accomplishments at 42% is more than twice that of achievements or activities. However, when the distribution is recalculated using the within-category analysis (the arrow points down, indicating that the total is calculated within each category), thus controlling for the difference in numbers of tokens produced: 80% of achievements, 75% of accomplishments, 52% of activities, and 40% of statives exhibit use of the preterite, supporting the AH. Similarly, the imperfect shows the greatest use with statives (60%), less with activities (48%), and approximately half of that with accomplishments (25%) and achievements (20%). The data from the English L1 learners is similarly presented (Table 2). Thus, the reanalysis challenges the authors' claim of counterevidence for the AH.

Table 2. Results of Dutch Learners by Vendler Category (across-category analysis; González & Quintana, 2018, Table 5, p. 620) and L1 English (Table 8, p. 621) and within-lexical-category reanalysis

	Achievements		Accomplishments		Activities		States		Total
	N	(%)	N	(%)	N	(%)	N	(%)	
Across lexical category									
L1 Dutch									
Preterite	43	(19)	96	(42)	44	(19)	46	(20)	229
Imperfect	11	(7)	32	(21)	40	(26)	70	(46)	158
Total	54		128		84		116		382
L1 Dutch reanalysis, Within lexical category									
Preterite	43	(80)	96	(75)	44	(52)	46	(40)	229
Imperfect	11	(20)	32	(25)	40	(48)	70	(60)	158
Total	54		128		84		116		382
Across lexical category									
L1 English									
Preterite	93	(32)	51	(18)	97	(34)	45	(16)	286
Imperfect	16	(19)	7	(8)	16	(19)	46	(54)	85
Total	109		56		113		91		371
L1 English reanalysis, Within lexical category									
Preterite	93	(85)	51	(88)	97	(86)	45	(49)	286
Imperfect	16	(15)	7	(12)	16	(14)	46	(51)	85
Total	109		56		113		91		371

In sum, the review of research addressing the AH has shown that the investigation of the acquisition of morphology tracking its emergence and development as it relates to the lexico-aspectual characteristics of verbal predicates has been a fruitful area of research, with past and ongoing debates. The picture that emerges from the review is that in its current formulation (i.e., Andersen, 2002) there are more studies whose results are consistent with the AH than not, and that a number of issues need to be examined, which is the goal of the following section.

The Aspect Hypothesis and SLA Research: Factors in the Use of Tense-Aspect

When authors reflect on explanations for variation in the acquisition of past morphology relative to the AH, four variables are frequently cited: proficiency, first language, input, and task (see Table 1 in the online supplementary materials). In this section we consider these variables through studies that have specifically set out to test their influence on the outcome of the AH.

We start with learner proficiency, which is directly addressed by the hypothesis, move to two issues related to learnability, namely language transfer and input, and conclude with investigations of task effect, which although crucial to research design, is not an acquisition issue, but rather a testing of acquisition issue. These variables (and more, including, e.g., discourse structure and narrative grounding) are in play simultaneously; however, we present them individually in the following sections.

Learner Proficiency

Testing learners at a range of proficiency levels is integral to testing the AH. Including beginning level learners is necessary to test the initial stages of morphological development following the prediction that "second language learners will *initially* be influenced by the inherent semantic aspect of verbs or predicates" (Andersen & Shirai, 1996, p. 533, emphasis added). In addition, studies need to test incrementally more proficient learners to observe the "spread" of past (preterite, imperfect, and progressive) across lexical aspectual categories to document not only initial but "subsequent uses of past forms" (Andersen, 2002, p. 79).

Researchers working with higher proficiency learners problematize the observation that the strength of the association of aspectual category and past morphology often increases with proficiency (e.g., Salaberry, 2008, p. 133; first noted by Robison, 1995). We do not find this problematic because the strengthening association is developmental and can be explained mathematically. Low-level learners who begin to use verbal morphology show very low levels of use of past morphology. They also use nonpast forms, such as base forms (or infinitives), present, and progressives (bare and inflected). As null morphology recedes, more verbs are inflected, and alternative forms become past forms (Bardovi-Harlig, 1998). For example, an analysis of oral narratives by a cross-section of learners of English (Bardovi-Harlig, 1998) shows that from the lowest group (Group 10-30, named for 10-30% overall use of past tense) to the next group (Group 50) base forms drop from 47% to 26% and alternative forms from 20% to 4%, and past rises from 32% to 71%. From Group 50 to 70, base forms drop to 11%, alternative forms stay at 4%, and past rises from 71% to 85%, and so on through the levels. In this manner, uninflected verbs become inflected and inflected verbs become past, and morphology associates with the lexical aspectual categories more strongly. Thus, increased association of past

morphology and aspectual category is a result of linguistic development in general, and it is difficult to consider it counterevidence to the AH.

Although there is currently broad interest in the advanced learner (cf. Vraciu, 2013), this does not always translate well into aspect studies. Several AH studies included learners that are too advanced to test the initial stages that Andersen (2002) described. More specifically, B2- and C2-level learners according to the Common European Framework of Reference for Languages, who have A-levels in French (the equivalent of high GPAs in high school majors in North America; McManus, 2013; see also Labeau, 2005) exhibit scores of greater than 80% use of perfective past (*passé composé*) in French (McManus, 2013, Table 8), which is not characteristic of beginning learners (cf. Bardovi-Harlig, 1998, Table 3, true beginners exhibit around 30% use of simple past in English). Proficiency labels have always been problematic in SLA, but using the term "low" for the lowest level of a relatively high-level group is counterproductive. Thus, if B2-C2 students are too proficient to exhibit early stages of perfective past, they seem to be in the middle of a trajectory toward the imperfect, which is acquired much later, demonstrating around 50% use of the *imparfait* in French (59% use of French imperfect with statives by year 1 university students, Labeau, 2005; 42% use of imperfect with atelics, McManus, 2013), which brings us to how more advanced learners can be instrumental in testing the AH.

Where intermediate-to-advanced learners have a genuine role to play is in the determination of when the imperfect enters the system and how it spreads. Research has documented that imperfect emerges after the perfective past, but has yet to determine the time interval and the effects of variables (such as instruction) on the spacing between perfective and imperfective past and the path that it travels. Camps (2005) and Comajoan (2006) captured the early stages of imperfect (for Spanish and Catalan, respectively) and provided a starting point for comparisons

to higher proficiency learners such as those in Izquierdo and Collins (2008) and McManus (2013). Furthermore, the performance of truly advanced learners can shed light on the use of non-prototypical associations, namely the preterite with statives and the imperfect with achievements (e.g., Diabubalick & Guijarro Fuentes, 2017; Domínguez et al., 2013; Izquierdo & Collins, 2008; Salaberry, 2002.)

Language Transfer

Two early longitudinal studies of L2 temporality concluded that there was a lack of noticeable first-language effects. Bardovi-Harlig (2000, p. 182) concluded that "there does not seem to be an obvious first language influence in the longitudinal corpus" produced by L1 speakers of Arabic, Japanese, Spanish, and Korean. Dietrich, Klein and Noyau (1995) observed that transfer "phenomena are strikingly rare. What is much more striking is the lack of SL [source language] influence where one would expect it," concluding that "there is no significant SL influence in the acquisition of temporality" (p. 278). They did not rule out transfer entirely, however, stating "We cannot exclude that clear transfer exists, of course; but if so, we have not observed it in the learner varieties studied in this project."

AH studies have subsequently investigated the effects of language transfer, using the hypothesis as a lens for identifying transfer. Although studies may control for task, proficiency, age, and educational levels (almost all participants are university students and all are instructed learners), studies differ in design, which can make comparisons difficult (Jarvis, 2000). Studies of a single L1 compare their results against other studies or against the predictions of the AH; other studies following Dietrich, Klein, and Noyau (1995) and Jarvis (2000) have compared two L1s (Table 2 in the online supplementary materials).

In this section we discuss studies that self-identified with the term *transfer* or *L1* in the research questions or title. More specifically, AH transfer studies ask a number of questions regarding whether a specific first language (dis)advantages the acquisition of a second-language tense-aspect system compared to one or more other languages:

(a) What is the influence of [+progressive] languages (English) and [–progressive] languages (German, Russian, Ukrainian, Bulgarian) on the acquisition L2 Japanese progressive and resultative? (Sugaya & Shirai, 2007)

(b) What is the influence of [+progressive] languages (Chinese, Japanese, Spanish) and [–progressive] languages (Hebrew, German, Polish) on the use of stative progressives in L2 English? (Fuchs & Werner, 2018)

(c) What is the influence of L1 French *passé composé* on English and use of present perfect in perfective (simple past) contexts (namely, achievements and accomplishments)? (Ayoun & Salaberry, 2008; Collins, 2002, 2004; Deshors, 2018)

(d) Is there an advantage of an L1 (or L2) that makes the perfect-imperfect distinction in acquiring an L2 (or L3) with the same or similar contrast? (Diaubalick & Guijarro-Fuentes, 2017; Fessi, 2016; Izquierdo, 2009; Izquierdo & Collins, 2008; Salaberry 2005)

(e) What is the influence of different tense-aspect systems in the L2 acquisition of tense-aspect? (L1 English, L1 Dutch on L2 Spanish, González & Quintana, 2018; L1 English, L1 German on L2 Italian, Giacalone Ramat, 2002; or L1 English, L1 German on L2 French, McManus, 2013)

These studies investigate whether the influence of transfer is sufficient to disrupt the sequences predicted by the AH, accelerating, slowing, or reordering development in some way. In this regard, Sugaya and Shirai (2007, p. 10) named two hypotheses addressing L1 transfer: a

strong L1 transfer hypothesis, in which L1 transfer could override the AH (Shirai, 2002b), and a weak L1 transfer hypothesis (which Sugaya & Shirai, 2007, attributed to Collins, 2002), in which L1 transfer does not override the AH, but rather transfer effects are found within the AH, such as limited use of the present perfect as a competitor to simple past in achievements and accomplishments in L2 English by L1 speakers of French (see question [c] above).

Several studies have investigated the influence of an L1 form that is structurally like a present perfect but that semantically is a perfective past on the acquisition of simple past or preterite. There have been enough such studies that they can be compared instructively: the L1s include French (Ayoun & Salaberry, 2008; Collins, 2002, 2004; Deshors, 2018) and Dutch (González & Quintana, 2018), the L2s English and Spanish; the studies focus on the use of present perfect in perfective contexts. The tasks have included cloze passages, essays, and narrative retells. Collins (2004) tested L1 French speakers (a language with a form equivalent to the English present perfect) and L1 Japanese speakers (a language without a present perfect form) using cloze passages. Both groups of learners exhibited use of simple past as predicted by the AH. However, both groups also used the perfect, contrary to the L1-based prediction. Although the L1 French learners primarily used the simple past, they also used the present perfect as an alternative to the simple past with telics. In contrast, the Japanese speakers were expected to not use the present perfect because it is not part of the Japanese tense-aspect system, but they nevertheless used present perfect and pluperfect as alternatives to the simple past in all categories, suggesting factors other than transfer may play a role. Ayoun and Salaberry (2008, p. 582) similarly reported that both cloze-passage and narrative data reflected a strong effect of lexical category, and that L1 French did not have a "pervasive negative influence" in that there were very few uses of the present perfect. Deshors (2018) examined the use of present perfect

with accomplishments and achievements (where the simple past is expected) in essays and critiques of literary texts from the French subsection of the *International Corpus of Learner English* and reported that the influence of the French *passé composé* was weak. Finally, González and Quintana (2018) compared L1 English and L1 Dutch learners of Spanish on a written film retell task in which only the Dutch learners used the present perfect in perfective past contexts (11% of the total); they also used the imperfect more than the English speakers who used preterite. González and Quintana concluded that both transfer and inherent aspect (dynamicity for L1 English and durative-terminative for L1 Dutch) determined the use of morphology using an across-category analysis.

Studies of Asian languages have also investigated the role of L1 transfer and the AH. Sugaya and Shirai (2007) investigated transfer and universals following Shirai (2002b) with two groups of learners, L1 progressive (English) and L1 nonprogressive (German, Russian, Ukrainian, Bulgarian) using two tasks. Results from the grammaticality judgment test showed that there was an effect of lexical aspect in how L2 learners of Japanese used progressive morphology (association of nonpast imperfective –*teiru* with progressive meaning with activity verbs, more common in the lower level; the past marker associated with achievements at the lower and higher levels). The data also showed that there was no L1 effect, since the L1 progressive and nonprogressive learners performed in similar ways.[10] Data from the oral picture description task investigating the use of progressive and resultative –*teiru* in the *present* exhibited an L1 effect in the lower level groups.[11] In sum, their results suggested that L1 transfer is evident only in the lower level, and it appears to operate within the AH.

In the aggregate, the studies show that L1 influence does not override lexical aspect, but rather, it is subtle and is mediated by two factors, proficiency and task. True low-level learners

show greater L1 influence than more advanced learners. This further supports the AH as a semantic and acquisitional universal, especially when we take into account the range of L1-L2 pairs that have been tested.

Language Input

The question of the influence of input on acquisitional patterns in tense-aspect acquisition was formalized by Andersen (1986, 1990) as the *distributional bias hypothesis.* Andersen and Shirai (1994) observed that "native speakers in normal interaction with other native speakers tend to use each verb morpheme with a specific class of verbs, also following the aspect hypothesis" (p. 137). Two types of studies have followed: analyses of large corpora that represent input that learners may encounter and studies of classroom input.

To characterize potential input to L2 learners of English, Wulff et al. (2009) investigated verb forms in the spoken section of the British National Corpus and the Michigan Corpus of Academic Spoken English (MICASE; Simpson et al., 2002). They concluded that frequency biases in the input result in certain verbs being more frequently associated with certain tense-aspect morphology (note that Wulff et al. discuss *verbs* not *predicates*). Verbs that are distinctively associated with particular tense-aspect forms (like activities with progressive) tend to be the prototypical associations. They concluded that "frequency alone is not enough" (p. 366) and that three factors—frequency, distinctiveness, and prototypicality—interact and are usually positively associated. Tracy-Ventura and Cuesta (2018) conducted a similar analysis using native-speaker corpus data for Spanish (*Corpus del español*, Davies, 2002). They found that the top 10 most frequent verbs associated with preterite in the native-speaker corpus were telic, whereas those with imperfect were atelic, and primarily states.

The second approach is to study the input available in specific learning environments (classroom input and instructor talk in classrooms). Camps (2005) illustrated the effect of classroom instruction by collecting narratives in response to the question "What was the best vacation you've ever had?" The first narratives were collected after instruction in Spanish preterite, but before instruction on the imperfect, and in these, the learners used only the preterite. Camps repeated the task after instruction on the imperfect, and learners used the imperfect as predicted by the AH (see "Review of Empirical Studies").

Whereas Camps described instructional input in terms of the forms taught and the length of instruction, later studies included an analysis of classroom language. Collins et al. (2012) collected 40 hours of classroom input from intensive EFL grade school classrooms in Quebec (6 classes of Grade 3), finding simple past in only 10% of the finite verbs forms in the corpus. 72% of past verbs appeared with telics (achievements/ accomplishments), 19% with activities, and 9% with statives. The retrospective verbal reports collected by Collins (2005) indicated that learners may be sensitive to not having encountered certain forms in input. Learners particularly commented on the lack of stative verbs in the past and their use in the present (p. 214), thus suggesting that similar input distributions may be found in a variety of classrooms.

Finally, Daidone (2019) examined input provided by teachers of Spanish. Recordings of instructor talk from twenty-four 50-minute classes of Spanish in third-, fourth-, and fifth-semester university courses (8 recordings per level) revealed that oral teacher input had more instances of preterite than imperfect (in a single class session an average of 23 tokens of preterite and 5 tokens of imperfect were provided). In fact, four instructors from the third and fourth semesters produced only the preterite (three were native speakers and only one was a non-native speaker of Spanish). She also compared teacher input to Spanish native-speaker corpus data from

Corpus de referencia del español actual (CREA) and the *Corpus del español* and found much greater use of imperfect in the corpus data than in teacher input.

These classroom-based studies raise the issue of impoverished input. The form-meaning associations are greatly simplified in at least some foreign language classrooms. If Spanish foreign-language classrooms are similar to those described in these studies, this would contribute to the robust use of preterite and the absence of imperfect reported by Salaberry (1999, 2000). An analogous situation for L2 Japanese is reported by Ishida (2004), in which the resultative reading of *-teiru* was taught exclusively 4 months before the progressive reading (2nd semester vs. 3rd semester), contributing to the resultative-before-progressive order, in contrast to the progressive-before-resultative order reported in all other studies (Shirai, 2002b). In contrast, pre-college ESL learners in academic programs in North America may encounter less controlled input in the academic environment and beyond, leading to the production a range of forms and meanings reported by Bardovi-Harlig (2000) and Collins (2002).

Task Effects

The effect of task is part of AH research for two reasons: a) to test generalizability, i.e., to investigate whether the predictions of the AH are borne out across tasks (there may also be an undercurrent of looking for a task in which the AH does not hold) and b) to facilitate data collection by providing production or interpretation opportunities that result in the greatest number of tokens in the greatest number of aspectual categories (see Table 3 in the online supplementary materials). Longitudinal studies tend to use a variety of tasks both to provide a wide range of tokens and to maintain participant interest (e.g., Bardovi-Harlig, 2012; Dietrich, Klein, & Noyau, 1995, data reanalyzed for the AH by Chan et al., 2012). Production tasks have

dominated interpretation and judgment tasks; the production tasks have predominately been written, with approximately half as many oral tasks, and half again eliciting both oral and written production.

In tense-aspect research, the gold-standard tasks have been those that allow learners to produce personal narratives. The studies reviewed here often treat narrative as a task. However, telling the narrative is the task and the resulting narrative is a *text*. Narratives are only one of a range of text-types, each with their own structure and tense-aspect use. Narratives are texts in which "the speaker relates a series of real or fictive events in the order in which they took place" (Dahl, 1984, p. 116). They are highly desirable because narrative foreground is sequenced, follows chronological order, and provides common contexts for the use of perfective past. Personal narratives are desirable because narrators are inclined to set the scenes for various events, provide motivations for actions, as well as interpretations and evaluations, all candidates for the use of imperfect in the narrative background.

Because learners in the initial stages of acquisition of tense-aspect morphology (the learner population addressed by the AH) are often unable to sustain lengthy personal narratives, and because personal narratives are not reliably comparable across learners (regardless of level of proficiency), researchers often elicit impersonal narratives. The use of impersonal narratives increases comparability, and gives every learner the same amount content to convey, if they are able to. However, this tends to sacrifice background information, and hence eliminates the environment for imperfective (Bardovi-Harlig, 2000, 2005; Noyau, 1990).[12] This has been supported further by Camps's (2005) comparison of personal narratives with imperfect to the impersonal narratives that lack imperfect in Salaberry (1999, 2000), and Comajoan's (2005) comparison of personal narratives in conversation and impersonal narratives in retellings,

resulting in more imperfect in conversational narratives than in retellings. Domínguez et al. (2013) addressed this issue by eliciting narratives in three tasks with increasing degrees of control over content: a) personal narratives in a guided interview, b) impersonal narratives in a picture retell task in which learners were provided with pictures and a habitual introduction (an imperfect context) "todas las mañanas eran las mismas" (*every morning was the same*) followed by a statement that introduced a sequence of events (a perfective context) "hasta que un día" (*until one day*); and c) a guided impersonal narrative in a picture narration task in which pictures were presented with 25 corresponding predicates and adverbial phrases below the pictures in a type of oral cloze test.

Before we leave narratives and the tasks that elicit them, we must consider that every narrative brings a potential confound that is largely unacknowledged by AH studies, namely narrative structure. Narrative discourse is comprised of two parts, the foreground and the background. The foreground relates events belonging to the skeletal structure of the discourse (Hopper, 1979) and consists of clauses that move time forward (Dry, 1981, 1983; see also Fleischman, 1985; Reinhart, 1984; von Stutterheim, 1991). In contrast, the background has many functions which together support the foreground. Grounding, like lexical aspect, influences the use of tense-aspect morphology (Hopper, 1979). Like achievements and accomplishments, activities that are sequenced in the foreground easily appear in the perfective past. Bardovi-Harlig (1998, 2000) demonstrated that grounding influences tense-aspect distribution, and that with the exception of achievements (which occur with perfective past regardless of grounding), accomplishments show higher use of perfective past than activities (as predicted by the AH), and that accomplishments and activities show higher use of perfective past in the foreground than in the background in learner narratives. Not surprisingly, background activities show robust use of

progressive, whereas foreground activities do not, instead appearing in the perfective past. While some studies acknowledge this variable in their literature reviews, only a few include grounding in the analysis (e.g., Comajoan, 2005; Domínguez et al., 2013; Howard, 2004; Labeau, 2005; Salaberry, 2011).

After telling narratives, the most common task in AH research is the cloze passage. Multiple cloze passages allow the greatest flexibility in balancing production across aspectual categories and contexts for perfective and imperfective morphology, which is harder to do naturally in a single narrative. Cloze passages privilege sufficient tokens for evaluation, comparability of production across learners, and balance of aspectual categories over communication (Bardovi-Harlig & Reynolds, 1995; Collins, 2002; Izquierdo, 2008; Izquierdo & Collins 2009; but see Ayoun & Salaberry, 2008; Zhao & Shirai, 2018, for a 59-item cloze that is not balanced). The potential use of explicit knowledge is minimized by asking the learners to work quickly. Domínguez et al. (2013) used guided impersonal narratives which were elicited by presenting sequenced pictures with verbs provided. In contrast to the open cloze passages which allow learners to use any form in their linguistic inventory (including base forms), forced-choice tasks between perfective and imperfective past (Salaberry, 2011) artificially elevate scores by giving learners only two target forms from which to choose.

Bardovi-Harlig (2000) found no difference in the degree of support for the AH from unconnected short cloze passages and either written or oral impersonal narratives; in the written narrative and cloze passage, achievements and accomplishments exhibited the highest rate of use of past morphology, followed by activities. In the oral narrative, achievements showed higher use of past than accomplishments, followed by activities. The rates of past tense use were highest in the written cloze passage (Bardovi-Harlig & Reynolds, 1995), next highest in the written

35

narrative, and lower in the oral narratives (Bardovi-Harlig, 1998). Collins (2002) replicated the cloze task with a different L1 and different learner population with the same results.

Shirai (2004) speculated that cloze passages may better support the AH than narratives, but in her review of the literature Bonilla (2013) found the opposite—that narratives better support the AH than cloze passages. This led to additional comparisons of narratives and cloze passages in French (oral and written *Modern Times* film retell and cloze passage; Labeau, 2005), Spanish (9-item cloze passage and one of either a personal narrative or a fairy tale; Ayoun & Salaberry, 2008), and English (the cloze passage, from Ayoun & Salaberry, 2008, and an oral personal narrative; Zhao & Shirai, 2018). Ayoun and Salaberry found that both the narrative and the cloze passage supported the AH. Zhao and Shirai reported that only the narrative did, and Labeau reported different analyses for the narrative and cloze passage and reported "bias provoked" by data collection methods, but the learners were likely too advanced for the AH to apply.

Cloze passages differ considerably across studies, most notably whether they represent connected texts or short independent passages. They may also differ within studies; Salaberry (2002) used three cloze passages from literary texts, and one he constructed to test nonprototypical combinations. In addition, all cloze passages differ from learner-constructed texts (of any type) because learners do not construct the cloze but rather supply verbs one at a time. Impersonal narratives also differ from each other in terms of events and subsequent encoding of those events and situations (Comajoan, 2006); similarly, Deshors (2018) noted that the different prompts used to elicit argumentative texts could have influenced their verb choices. This variation emphasizes the fact that tests of task effects occur in the context of both varying learner proficiency and task construction.

Among the studies that have identified task as a central concern, one nevertheless finds problematic procedures often caused by the pressure to elicit sufficient tokens to test the hypothesis. Open-book elicitation, in which the learners look at the pictures during the task, increases deixis, including the use of non-past and progressive, thus sacrificing the test of past required by the AH (Muñoz & Gilabert, 2018). Directions that tell learners to use past tenses (González & Quintana, 2018) or to write at least 15 verbs and 4 adverbs (Ayoun & Salaberry, 2008) may encourage the use of explicit knowledge on the part of the learners. Giving learners two days in advance to prepare a personal narrative or a fairy tale assures that learners can complete the task, but it also invites planning and rehearsal even when learners were told not to write their stories (Ayoun & Salaberry, 2008). Instructions may also direct learners away from the past as in "Describe what you've just seen: what are the characters doing?" (Rastelli & Vernice, 2013). In this case, the authors claim that the learners were free to use either past or present, but the directions prime the learners for nonnarrative by using "describe" and present progressive "What *are* the characters *doing*?" Also problematic are studies that do not report how the task was performed (e.g., McManus, 2013).[13]

Researchers can avoid the pitfalls of over explicit instructions by eliciting texts through defining questions. Von Stutterheim and Klein and colleagues have suggested that texts can be distinguished by the question that they answer, known as the *Quaestio* (von Stutterheim, 1991; Carroll & von Stutterhiem, 2003; von Stutterheim & Klein, 1989; von Stutterheim, & Lambert, 2005). This provides a practical solution to data collection because researchers can use such questions to elicit texts: Narratives answer the question, "What happened next?" (von Stutterheim & Klein, 1989), descriptions "What did/does X look like?" (von Stutterheim & Klein, 1989) or "What was it like there?" (Carroll & von Stutterheim, 2003), and arguments

"What is the evidence for your position?" (Bardovi-Harlig, 2013). Nonnarratives are particularly good linguistic contexts for imperfects and are often the type of text collected in corpora (e.g., Deshors, 2018).

In sum, the AH studies reviewed here show that the preponderance of tasks elicit narratives; however, they also show that narrative texts are often not distinguished from the tasks that elicit them. Eliciting nonnarrative text types, such as descriptions and argumentative texts, through appropriate tasks would provide evidence of the acquisition of other forms (e.g., imperfect, progressive). Tasks that elicit texts are one way to study the use of tense-aspect morphology in relatively lengthy learner-constructed language samples; conversations and conversational interviews are another. Production tasks dominate research addressing the AH in the last 20 years. As we become increasingly interested in non-prototypical associations, we may include truth-value tasks and acceptability judgment tasks to further probe form-meaning associations.

Moving Forward with Research on the Aspect Hypothesis

This section positions the AH in the larger inquiry of SLA, considering the role of both description and explanation with respect to the hypothesis and it identifies areas for future research within the domain of the AH that will expand our knowledge of how L2 tense-aspect systems are acquired.

Some criticisms of research in the AH have raised the issue of the explanatory power of the hypothesis, in the sense that it may provide an accurate description of interlanguage development but falls short of an explanation (Labeau, 2005; Salaberry, 2008; Slabakova, 2002; see also Andersen, 2002). In this respect, it may be useful to make a distinction between theory and hypothesis (VanPatten & Williams, 2015, p. 5):

Distinct from a theory, a hypothesis does not unify various phenomena; it is usually an idea about a single phenomenon. Some people use theory and hypothesis interchangeably, but in fact, they are distinct and should be kept separate. In science, we would say that a theory can generate hypotheses that can then be tested by experimentation or observation.

Following VanPatten and Williams' definition, we can say that the AH does not unify a number of phenomena related to the acquisition of tense-aspect, that is, everything related to the acquisition of tense-aspect in all contexts and tasks and from all L1s. Instead, the AH is a prediction about one phenomenon, namely the early association of past morphology, both perfective and imperfective (and within that both general imperfective and progressive) with lexico-semantic categories of verbal predicates. The explanation for this attraction is the underlying semantics. The encompassing theory in which the AH was proposed is functionalism, and within functionalism, the AH is just one piece of a larger puzzle that is the use of past tense morphology within discourse (or task, when eliciting data).

In his discussion of SLA theories, Jordan (2004) distinguishes between the objectives and the domain of theories. *Objectives* refers to whether theories have goals that are more descriptive (making sense of data) or explanatory (finding causation). Extending this approach to hypotheses, the AH provides both description and explanation: It provides a thorough description of learner data regarding tense-aspect and it provides an explanation for it, namely, the lexical semantics of the verbal predicates influence learners to inflect them with certain morphemes earlier or later. However, as we have shown in the review, this pattern is mediated by additional variables, which should be accounted for, not by an expanded AH, but by other interacting components of a theory of SLA. Jordan defines the domain of a theory of SLA as the number of

39

phenomena and their boundaries that the theory should attempt to explain (e.g., learner internal vs. learner external phenomena). The domains of the different perspectives on the acquisition of tense-aspect morphology are heterogeneous, but studies discussing the AH mostly attempt to explain the same domain, namely the acquisition of verb morphology in relation to lexical aspect. In this respect, one cannot ask from the AH a complete account of L2 acquisition, because the AH at its core makes linguistic claims that will necessarily interact with other learning variables (e.g., discourse grounding in narratives).

In this view, not all tense-aspect phenomena are of immediate relevance to the AH and thus not all of them require an explanation from the AH. For instance, one factor that has been almost assiduously ignored in research in the AH is the influence of narrative structure (Bardovi-Harlig, 2000, and Salaberry, 2011, are two among few exceptions). We do not suggest here that every test of the AH also be a test of the discourse hypothesis, but rather that studies using narratives cease to ignore the internal structure of narratives (namely, foreground and background) that also exerts influence on tense-aspect morphology. By extension, understanding the relation of text types to tense-aspect use through both the text type (Smith, 2003) and main and side structures within those text types (von Stutterheim, 1991; Carroll & von Stutterhiem, 2003; von Stutterheim & Klein, 1989; von Stutterheim, & Lambert, 2005) will allow researchers to explore imperfective and nonnarrative text types.

In sum, in contrast to calls for extending the AH to include additional variables (e.g., Labeau, 2005), we advocate keeping the relevant hypotheses separate. Theories of SLA need accounts of transfer, input, instruction, and task effects, but those accounts are separate from the linguistically motivated hypotheses concerning the development of specific constructs. A theory of language transfer should apply to every relevant case, not just tense-aspect. Similarly,

hypotheses concerning the effects of tasks should be stated generally, and apply to all relevant cases; the same with input and a special case of input, instruction.

As important as it is to understand what is beyond the scope of the AH, it is also important to understand what type of continued research in the domain of the AH would be hypothesis-driven and also expand our knowledge of L2 tense-aspect systems. We suggest three areas of investigation here.

a) What is the timing of the emergence of preterite (or perfective past) and the imperfect for languages that maintain that distinction? One area of investigation that arises from unresolved differences in the empirical data is the timing of the acquisition of the perfective and imperfective pasts. In his schematic representation of the stages of acquisition (see Figure 1), Andersen (1991) seems to have proposed rather rapid acquisition of imperfect, in stage 3, following the emergence of perfective past in stage 2, out of 8 stages. Similarly, Domínguez et al. (2013) referred explicitly to the rapid acquisition of the imperfect in their rephrasing of the predictions of the AH: "The imperfective marker appears soon after the perfective marker is first used" (p. 561, prediction 3). In contrast, Andersen (2002, p. 83) revised his estimate, instead claiming that the imperfective does not appear "for a long time," even though the perfective has begun to spread across categories. Bardovi-Harlig (2005) read the empirical literature to suggest, supported by the additional studies reviewed here, that the imperfect takes longer to emerge than Andersen (1991) had hypothesized. These differences in the timing of the emergence of imperfect morphology are evidence that perfective and imperfective morphology may develop at different rates and that even though lexical aspect influences the development of both, they are not necessarily timed as first posited by Andersen (1991).

b) What is the relationship between the general imperfective and the progressive in languages that exhibit both? (Hypotheses 2 and 3). Although Hypotheses 1 (the spread of perfective past) and 2 (the spread of imperfect past) are frequently investigated together, Hypotheses 2 and 3 (the spread of the past progressive) have not been. Whereas the perfective and imperfective have a hypothesized order of emergence, no such order has been investigated for general imperfectives and the progressive in languages that have both, such as Spanish, Catalan, Portuguese, and Italian. Addressing such a question would also entail determining the relationship between the semantics of the general imperfective (which includes habituality, progressivity, and continuity or unboundedness) and the progressive (see Giacalone Ramat, 1997).

c) Of the three areas that we suggest for further research in the framework of the AH, the most intriguing to us is the investigation of nonprototypical associations of past morphology and lexical aspectual categories, an area which leads back to the semantics of aspect. A number of research questions could be further investigated, such as: (i) Do learners produce such associations, (ii) Do learners understand nonprototypical associations; that is, can they compute the coercions? (iii) If so, when are they able to do so? As we suggested in earlier sections, the comprehension and use of nonprototypical associations of past morphology and lexical aspectual categories is very likely the domain of advanced learners, but this too is an empirical question. Narratives will show predominately prototypical associations, and nonprototypical uses will be relatively rare. Production studies will need to create contexts for four-way contrasts (telic/atelic x perfective/imperfective), creating four combinations (two prototypical and two nonprototypical), whereas interpretation tasks will permit the testing of more nuanced cases.

We identified four studies that have begun to address this question (Salaberry 2002; Izquierdo & Collins, 2008; Domínguez et al., 2013; Diaubalick & Guijarro Fuentes, 2017). As is the case with new areas of study, there is little consistency in the definition of "nonprototypical associations" in these studies. For instance, Salaberry (2002, p. 401) considered that "The non-prototypical value is represented by the case when the aspectual inflection is in direct contradiction with lexical aspect" (e.g., use of imperfect in telic predicates). Some studies have investigated the use of imperfect in telic predicates and perfective in atelic predicates (Izquierdo & Collins, 2008; Diaubalick & Guijarro Fuentes, 2017 with external adverbials), and another has focused on grounding, investigating prototypical combinations such as states in the background and achievements in the foreground, and nonprototypical combinations of states in foreground, and achievements in the background (Domínguez et al., 2013). However, achievements are not unusual in the background. The only class that seems to be at all restricted are statives in the foreground because they cannot be sequenced. In this respect, such contexts do not test non-prototypicality in relationship to the AH, but rather in interaction with the discourse hypothesis, and thus, the learners' use of morphology due to narrative grounding is difficult to untangle from the influence of lexical aspect. Further research on the acquisition of non-prototypical uses of tense-aspect morphology should explicitly provide a definition of prototypicality and how it applies to lexical aspect and/or discourse grounding as well as other features, such as adverbials.

Following Smith (1997), we suggest that nonprototypical associations used to test the AH consist of combinations of lexical aspect features and verbal morphology that require reinterpretation of a situation, and that these would ideally be tested on the predicate and tense-aspect morphology in isolation. For example, in the imperfective of achievements, the imperfective presents the preliminary stages of the event. In sentences like "The team was

reaching the top"/ "El equipo estaba alcanzando la cima" or "She was winning the race"/ "La corredora estaba ganando la carrera," there is no suggestion that the achievement actually takes place. Thus, in a yes/no interpretation task, given the sentence, "The team was reaching the top," the answer to *Did the team reach the top?* would be "no" or "I don't know" and the answer to *Did the team climb on the mountain?* would be "yes." In a truth value task, given the sentence "She was winning the race," the sentence "She was ahead of other runners" would be true, but "She won the race" would be false or unknown.

Similarly, statives with progressive are generally not allowed (such as "*Bill was knowing the answer"), but with "be" statives allow reinterpretation as in "John was being a hero" (Smith, 1997). With "be" statives, the progressive usually indicates that the states are transitory "stage-level ways of behaving that involve will and energy, although not necessarily overt activity" (p. 58). These sentences could be tested by truth value tasks in which the target sentence and one interpretation (at a time) is presented (see also Slabakova, 2003).

1. John was being a jerk.   2. Jane was being quiet

a. John was usually a jerk (false)   a. Jane was usually quiet (false)

b. John was not usually a jerk (true)   b. Jane was not usually quiet (true)

Accomplishments with progressive are not read as completed, which could be tested in a judgment task:

3a. Mary was walking to school, but she didn't actually get there (good)

b. #Mary walked to school, but she didn't actually get there (not good)

Nonprototypicality within the AH is a type of theoretically defined difficulty, but only a subset of what might constitute interesting learnability issues in the domain of tense-aspect systems. Other areas of "difficulty" may be caused, for example, by the many readings of the

imperfect or other tense-aspect forms (such as present perfect vs. preterite readings), questions outside the scope of the AH.

Conclusion

This paper has reviewed research of the past 20 years that has been conducted to investigate the AH. As we have shown, some studies speak directly to the AH, both in terms of theoretical constructs and research design, and others raise concerns either outside the AH or use designs or analyses that cannot be evaluated as claimed. The AH accurately predicts the adult L2 acquisition of past morphology in a number of languages. This is not to deny that a number of challenges related to proficiency level, L1 transfer, task variation, and methodology persist. However, variation in design, implementation, and analysis seems to confound rather than clarify the interpretation of results. For researchers to be able to compare results across studies, designs should be clearly stated (with tasks included in appendices or supplementary materials, or made available through IRIS), tests of lexico-aspectual categories and problematic cases identified and reported (Shirai, 2013), examples given, and total counts of data for analyses provided. More than 30 years after Andersen (1986) first posited the hypothesis, it is clear that it has generated a wealth of research and that it continues to inspire researchers to investigate the link between lexico-aspectual information and the acquisition of L2 morphology. We have outlined three areas of further investigation within the AH as well as areas beyond the scope of the hypothesis. We hope that researchers will pursue both lines of investigation for a better understanding of the acquisition of L2 temporal systems.

References

Andersen, R. W. (1986). El desarollo de la morfología verbal en el Español como segundo idioma [The development of verbal morphology in Spanish as a second language]. In J. M. Meisel (Ed.), *Adquisición de lenguaje/ Aquisição da linguagem [Acquisition of language]* (pp. 115–138). Vervuert.

Andersen, R. W. (1990). Models, processes, principles and strategies: Second language acquisition inside and outside the classroom. In B. VanPatten & J. F. Lee (Eds.), *Second language acquisition—foreign language learning* (pp. 45–78). Multilingual Matters.

Andersen, R. W. (1991). Developmental sequences: The emergence of aspect marking in second language acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 305–324). John Benjamins.

Andersen, R. W. (2002). Dimensions of "pastness." In R. Salaberry & Y. Shirai (Eds.), *Tense-aspect morphology in L2 acquisition* (pp. 79-105). John Benjamins.

Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition, 16*, 133–156.

Andersen, R. W., & Shirai, Y. (1996). The primacy of aspect in first and second language acquisition: The pidgin-creole connection. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 527–570). Academic Press.

Antinucci, F., & Miller, R. (1976). How children talk about what happened. *Journal of Child Language, 3*, 167–189

Ayoun, D., & Salaberry, M. R. (2008). Acquisition of English tense-aspect morphology by advanced French instructed learners. *Language Learning, 58,* 555–595.

Bardovi-Harlig, K. (1992). The use of adverbials and natural order in the development of

    temporal expression. *IRAL, 30*, 299–320.

Bardovi-Harlig, K. (1994). Anecdote or evidence? Evaluating support for hypotheses concerning

    the development of tense and aspect. In E. Tarone, S. M. Gass, & A. D. Cohen (Eds.),

    Research methodology in second language acquisition (pp. 41–60). Erlbaum.

Bardovi-Harlig, K. (1998). Narrative structure and lexical aspect: Conspiring factors in second

    language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition*,

    20, 471-508.

Bardovi-Harlig, K. (1999). From morpheme studies to temporal semantics: Tense-aspect

    research in SLA. *Studies in Second Language Acquisition, 21*, 341-382.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning,*

    and use. Blackwell.

Bardovi-Harlig, K. (2002). Analyzing aspect. In R. Salaberry & Y. Shirai (Eds.), *Tense-aspect*

    morphology in L2 acquisition (pp. 129-154). John Benjamins.

Bardovi-Harlig, K. (2005). Tracking the elusive imperfect in adult second language acquisition:

    Refining the hunt. In P. Kempchinsky & R. Slabakova (Eds.), *Aspectual inquiries* (pp. 397-

    419). Kluwer.

Bardovi-Harlig, K. (2012). After process, then what? A longitudinal investigation of the

    progressive prototype in L2 English**.** In E. Labeau & I. Saddour (Eds.), *Development of*

    tense, aspect and mood in L1 and L2 (pp. 135-155). Rodopi.

Bardovi-Harlig, K. (2013). Research design: From text to task. In R. Salaberry & L. Comajoan

    (Eds.) *Research design and methodology in studies on second language tense and aspect*

    (pp. 219-269)*.* Mouton de Gruyter.

Bardovi-Harlig, K., & Reynolds, D. W. (1995). The role of lexical aspect in the acquisition of

    tense and aspect. *TESOL Quarterly, 29*, 107-131.

Binnick, R. I. (1991). *Time and the verb: A guide to tense and aspect*. Oxford University Press.

Binnick, R. I. (Ed.) (2012). *The handbook of tense and aspect.* Oxford University Press.

Bloom, L., Lifter, K., & Hafitz, J. (1980). Semantics of verbs and development of verb inflection

    in child language. *Language, 56*, 386–412.

Bonilla, C. L. (2013). Tense or aspect?: A review of initial past tense marking and task

    conditions for beginning classroom learners of Spanish. *Hispania, 96,* 624-639.

Bronckart, J.-P., & Sinclair, H. (1973). Time, tense and aspect. *Cognition, 2,* 107-130.

Camps, J. (2005). The emergence of the imperfect in Spanish as a foreign language: The

    association between imperfective a morphology and state verbs. *IRAL, 43,*163-192.

Carroll, M. & von Stutterhiem, C. (2003). Typology and information organisation: Perspective

    taking and language-specific effects in the construal of events. In A. Giacalone Ramat (Ed.),

    Typology and second language acquisition (pp. 365-402*)*. Mouton de Gruyter.

Chan, H. L., Finberg, J., Costello, W., & Shirai, Y. (2012). L2 acquisition of tense-aspect

    morphology: Lexical aspect, morphological regularity, and transfer. In L. Filipović & K.

    Jaszczolt (Eds.), *Space and time in languages and cultures* (pp. 181-204). John Benjamins.

Collins, L. (2002). The roles of L1 influence and lexical aspect in the acquisition of temporal

    morphology. *Language Learning, 52*, 43-94.

Collins, L. (2004). The particulars on universals: A comparison of the acquisition of tense-aspect

    morphology among Japanese- and French-speaking learners of English. *The Canadian

    Modern Language Review, 61*, 251-274.

Collins, L. (2005). Accessing second language learners' understanding of temporal morphology. *Language Awareness, 14,* 207-220.

Collins, L., White, J., Trofimovich, P., Cardoso, W., & Horst, M. (2012). When comprehensible input is not comprehensive input: A multi-dimensional analysis of past tense in classroom input. In C. Muñoz (Ed.) *Intensive exposure experiences in second language learning* (pp. 66-87). Multilingual Matters.

Comajoan, L. (2005). The emergence of L2 imperfective morphology in narrative retellings and conversations. *Catalan Review, 19*, 79-108.

Comajoan, L. (2006). The aspect hypothesis: Development of morphology and appropriateness of use. *Language Learning, 56*, 201–268.

Comajoan, L., & Pérez-Saldanya, M. (2005). Grammaticalization and language acquisition: Interaction of lexical aspect and discourse. In D. Eddington (Ed.), *Selected proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 44-55). Cascadilla Proceedings Project.

Comrie, B. (1976). *Aspect*. Cambridge University Press.

Comrie, B. (1985). *Tense*. Cambridge University Press.

Dahl, Ö. (1984). Temporal distance: Remoteness distinctions in tense-aspect systems. In B. Butterworth, B. Comrie, & Ö. Dahl (Eds.), *Explanations for language universals* (pp. 105–122). Mouton.

Daidone, D. (2019). Preterite and imperfect in Spanish instructor oral input and Spanish language corpora. *Hispania*, *102,* 45-58.

Davies, M. (2002). *Corpus del español*. https://www.corpusdelespanol.org/

De Swart, H. (2012). Verbal aspect. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 752-780). Oxford University Press.

Deo, A. (2012). Morphology. In R. Binnick (Ed.), *The Oxford handbook of tense and aspect* (pp. 155-183). Oxford University Press.

Deshors, S. C. (2018). Does the passé composé influence L2 learners' use of English past tenses? A semantic exploration of the present perfect in French-English interlanguage. *International Journal of Learner Corpus Research, 4*, 23–53.

Diaubalick, T. & Guijarro Fuentes, P. (2017). L1 effects as manifestations of individual differences in the L2 acquisition of the Spanish tense-aspect-system. In K. Bellamy, M. W. Child, P. González, & A. Muntendam (Eds.), *Multidisciplinary approaches to bilingualism in the Hispanic and Lusophone world* (pp. 9-40). John Benjamins.

Dietrich, R., Klein,W., & Noyau, C. (1995). *The acquisition of temporality in a second language.* John Benjamins.

Domínguez, L., Tracy-Ventura, N., Arche, M. J., Mitchell, R., & Myles, F. (2013). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition, 16,* 558–577.

Dowty, D. (1979). *Word meaning and Montague grammar.* Reidel.

Dry, H. (1981). Sentence aspect and the movement of narrative time. *Text, 1*, 233-40.

Dry, H. (1983). The movement of narrative time. *Journal of Literary Semantics, 12*, 19-53.

Ellis, R. (2018). Meta-analysis in second language acquisition research. A critical appraisal. *Journal of Second Language Studies, 1*, 231-253.

EPPI-Centre. (2006). *EPPI-Centre methods for conducting systematic reviews*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Fessi, I. (2016). Cross-linguistic influence in tense-aspect Spanish L3 acquisition: A study of

    Arabic Tunisian learners of L3 Spanish. *Revista Nebrija de Lingüística Aplicada, 20*.

Fleischman, S. (1985). Discourse functions of tense-aspect oppositions in narrative: Toward a

    theory of grounding. *Linguistics, 23*, 851-882.

Fuchs, R., & Werner, V. (2018). The use of stative progressives by school-age learners of

    English and the importance of the variable context: Myth vs. (corpus) reality. *International

    Journal of Learner Corpus Research, 4*, 195–224.

Giacalone Ramat, A. (1997). Progressive periphrases, markedness and second language data. In

    S. Elliasson & E. H. Jahr (Eds.), *Language and its ecology: Essays in memory of Einar

    Haugen* (pp. 261-285). Mouton de Gruyter.

Giacalone Ramat, A. (2002). How do learners acquire the classical three categories of

    temporality? In R. Salaberry & Y. Shirai (Eds.), *Tense-aspect morphology in L2 acquisition*

    (pp. 221-247). John Benjamins.

González. P., & Quintana Hernández, L. (2018). Inherent aspect and L1 transfer in the L2

    acquisition of Spanish grammatical aspect. *The Modern Language Journal, 102,* 611-625.

Hopper, P. J. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Syntax and

    semantics: Discourse and syntax* (pp. 213-241). Academic Press.

Howard, M. (2004). On the interactional effect of linguistic constraints on interlanguage

    variation: The case of past time marking, *IRAL, 42,* 319-334.

Ishida, M. (2004). The effect of recasts on acquisition of aspect in Japanese as a foreign

    language. *Language Learning*, *54*, 311–394.

Izquierdo, J. (2009). L'aspect lexical et le développement du passé composé et de l'imparfait en

    français L2: une étude quantitative auprès d'apprenants hispanophones. *The Canadian*

    Modern Language Review, 65, 587-613.

Izquierdo, J., & Collins, L. (2008). The facilitative role of L1 influence in tense–aspect marking:

    A comparison of Hispanophone and Anglophone learners of French. *The Modern Language*

    Journal, 92, 350-368.

Izquierdo, J., & Kihlstedt, M. (2019). L2 imperfective functions with verb types in written

    narratives: A cross-sectional study with instructed Hispanophone learners of French. *The*

    Modern Language Journal, 103, 291-307.

Jarvis, S. (2000). Methodological rigor in the study of transfer: Identifying L1 influence

    in the interlanguage lexicon. *Language Learning, 50,* 245-309.

Kaplan, M. A. (1987). Developmental patterns of past tense acquisition among foreign language

    learners of French. In B. VanPatten, T. R. Dvorak, & J. F. Lee (Eds.), *Foreign language*

    learning: A research perspective (pp. 52–60). Newbury House.

Kenny, A. (1963). *Action, emotion, and will.* Routledge and Kegan Paul.

Kim, H.-J. (2012). A case study of tense-aspect marking by L2 learners of Korean. In E. Labeau

    & I. Saddour (Eds.), *Tense, aspect and mood in first and second language acquisition* (pp.

    153-177). Rodopi/Cahiers Chronos.

Klein, W. (2009). How time is encoded. In W. Klein & L. Ping (Eds.), *The expression of time*

    (pp. 39–82). Mouton de Gruyter.

Labeau, E. (2005). Beyond the aspect hypothesis: Tense-aspect development in advanced L2

    French. *EuroSLA Yearbook, 5,* 77-101.

Lardiere, D. (2003). Revisiting the comparative fallacy: A reply to Lakshmanan and Selinker

2001. *Second Language Research, 19*, 129–143

Lee, E. H., & Kim, H. Y. (2007). On cross-linguistic variations in imperfective aspect: The case of L2 Korean. *Language Learning, 57,* 651–685.

Li, P., & Shirai, Y. (2000). *The acquisition of lexical and grammatical aspect*. Mouton de Gruyter.

McManus, K. (2013). Prototypical influence in second language acquisition: What now for the Aspect Hypothesis? *IRAL, 51,* 299–322.

Muñoz, C., & Gilabert, R. (2011). More evidence concerning the Aspect Hypothesis: The acquisition of English progressive aspect by Catalan-Spanish instructed learners. *IRAL, 49*, 241–263.

Noyau, C. (1990). The development of means for temporality in the unguided acquisition of L2: Cross-linguistic perspectives. In H. W. Dechert (Ed.), *Current trends in European second language acquisition research* (pp. 143-170). Multilingual Matters.

Rastelli, S. (2019). The imperfective paradox in a second language: A dynamic completion entailment test. *Lingua, 231*, 1-23.

Rastelli, S., & Vernice, M. (2013). Developing actional competence and the building blocks of telicity in L2 Italian. *IRAL, 51*, 55–75.

Reinhart, T. (1984). Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics, 22*, 779-809.

Robison, R. E. (1990). The primacy of aspect: Aspectual marking in English interlanguage. *Studies in Second Language Acquisition, 12*, 315–330.

Robison, R. E. (1995). The aspect hypothesis revisited: A cross-sectional study of tense and aspect marking in interlanguage. *Applied Linguistics, 16*, 344–370.

Rohde, A. (1996). The aspect hypothesis and emergence of tense distinction in naturalistic L2 acquisition. *Linguistics, 34,* 1115–1138.

Ryu, J-Y., Horie, K., & Shirai, Y. (2015). Acquisition of the Korean imperfective aspect markers *–ko iss–* and *–a iss–* by Japanese learners: A multiple-factor account. *Language Learning, 65,* 791–823.

Salaberry, M. R. (1999). The development of past tense verbal morphology in classroom L2 Spanish. *Applied Linguistics, 20*, 151–78.

Salaberry, M. R. (2000). *The development of past tense morphology in L2 Spanish*. John Benjamins.

Salaberry, M. R. (2002). Tense and aspect in the selection of Spanish verbal morphology. In R. Salaberry & Y. Shirai (Eds.), *Tense-aspect morphology in L2 acquisition* (pp. 371-415). John Benjamins.

Salaberry, M. R. (2005). Evidence for transfer of knowledge of aspect from L2 Spanish to L3 Portuguese. In D. Ayoun & M. R. Salaberry (Eds.) *Tense and aspect in Romance languages: Theoretical and applied perspectives* (pp. 179-210). John Benjamins.

Salaberry, M. R. (2008). *Marking past tense in second language acquisition*. Continuum.

Salaberry, M. R. (2011). Assessing the effect of lexical aspect and grounding on the acquisition of L2 Spanish preterite and imperfect among L1 English speakers. *Bilingualism: Language and Cognition, 14*, 184–202.

Salaberry, M. R. & Comajoan, L. (Eds.) (2013). *Research design and methodology in studies on second language tense and aspect.* Mouton de Gruyter.

Salaberry, M. R., & Y. Shirai, Y. (Eds.). (2002). *The L2 acquisition of tense-aspect morphology.* John Benjamins.

Shirai, Y. (1991). *Primacy of aspect in language acquisition: Simplified input and prototype.* Unpublished doctoral dissertation, University of California, Los Angeles.

Shirai, Y. (2002a). The prototype hypothesis of tense-aspect acquisition in second language. In M. R. Salaberry & Y. Shirai (Eds.), *The L2 acquisition of tense-aspect morphology* (pp. 455-478). John Benjamins.

Shirai, Y. (2002b). The aspect hypothesis in SLA and the acquisition of Japanese. *Acquisition of Japanese as a Second Language*, *5*, 42–61.

Shirai, Y. (2004). A multiple-factor account for the form-meaning connections in the acquisition of tense-aspect morphology. In B. VanPatten, J. Williams, S. Rott, & M. Overstreet (Eds.), *Form meaning connections in second language acquisition* (pp. 91–112). Erlbaum.

Shirai, Y. (2007). The Aspect Hypothesis, the comparative fallacy and the validity of obligatory context analysis: A reply to Lardiere. *Second Language Research, 23*, 51–64.

Shirai, Y. (2013). Defining and coding data: Lexical aspect in L2 studies. In M. R. Salaberry & L. Comajoan (Eds.). *Research design and methodology in studies on L2 tense and aspect* (pp. 271-308). Mouton de Gruyter.

Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan corpus of academic spoken English.* Ann Arbor, MI: The Regents of the University of Michigan. *quod.lib.umich.edu/m/micase/*

Slabakova, R. (2002). Recent research on the acquisition of aspect: an embarrassment of riches? *Second Language Research, 18*, 172-188.

Slabakova, R. (2003). Semantic evidence for functional categories in interlanguage grammars. *Second Language Research, 19*, 42-75.

Smith, C. S. (1983). A theory of aspectual class. *Language, 59,* 479–501.

Smith, C. (1997). *The parameter of aspect* (2nd edition). Kluwer.

Smith, C. (2003). *Modes of discourse: The local structure of texts.* Cambridge University Press.

Sugaya, N. & Shirai, Y. (2007). The acquisition of resultative and progressive meanings of the imperfective aspect marker by L2 learners of Japanese: Transfer, universals, or multiple factors. *Studies in Second Language Acquisition, 29*, 1–38.

Tong, X., & Shirai, Y. (2016). L2 acquisition of Mandarin *zai* and *–le. Chinese as a Second Language Research, 5,* 1-25.

Tracy-Ventura, N., & Cuesta Medina, J. A. (2018). Can native-speaker corpora help explain L2 acquisition of tense and aspect? A study of the "input." *International Journal of Learner Corpus Research, 4,* 277-300.

Uno, M. (2014). Lexical aspect in the use of the present perfect by Japanese EFL learners. *IRAL, 52,* 31-57.

VanPatten, B., & Williams, J. (Eds.). (2015). *Theories in second language acquisition* (2nd edition). Erlbaum.

Verkuyl, H. J. (1993). *A theory of aspectuality: The interaction between temporal and atemporal structure*. Cambridge University Press.

Vendler, Z. (1967). Verbs and times. In Z. Vendler (Ed.), *Linguistics and philosophy* (pp. 97–121). Cornell University Press. (Reprinted from *Philosophical Review*, 1957, *66*, 143–160).

von Stutterheim, C., & Klein, W. (1989). Referential movement in descriptive and narrative discourse. In R. Dietrich & C. F. Graumann (Eds.), *Language processing in social context,* (pp. 39–76). Elsevier.

von Stutterheim, C. (1991). Narrative and description: Temporal reference in second language

    acquisition. In T. Huebner & C. A. Ferguson (Eds.), *Crosscurrents in second language*

    acquisition and linguistic theories (pp. 385-403). John Benjamins.

von Stutterheim, C., & Lambert, M. (2005). Cross-linguistic analysis of temporal perspectives in

    text production. In H. Hendricks (Ed.), *The structure of learner varieties* (pp. 203-230).

    Mouton de Gruyter.

Vraciu, A. (2013). Exploring the upper limits of the Aspect Hypothesis: Tense-aspect

    morphology in the advanced English L2 variety. *Language, Interaction and Acquisition, 4*,

    256–290.

Weist, R. M., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E., & Konieczna, E. (1984).

    The defective tense hypothesis: On the emergence of tense and aspect in child Polish.

    Journal of Child Language, 11, 347–374.

Wulff, S., Ellis, N. C., Römer, U., Bardovi-Harlig, K., & LeBlanc, C. (2009). The acquisition of

    tense-aspect: Converging evidence from corpora, and telicity ratings. *The Modern Language*

    Journal, 93, 354-369.

Zhao, H., & Shirai, Y. (2018). Arabic learners' acquisition of English past tense morphology:

    Lexical aspect and phonological saliency. *International Journal of Learner Corpus*

    Research, 4, 253–276.

Notes

[1] See Klein (2009) and Binnick (2012) for full descriptions of the concept of time in language.

[2] See Bardovi-Harlig (2000), Li and Shirai (2000), and Salaberry (2000, 2008) for a fuller introduction to lexical aspect in SLA.

[3] See Antinucci & Miller (1976); Bloom, Lifter, & Hafitz (1980); Bronckart & Sinclair (1973), for first language acquisition; Bardovi-Harlig (2000) for an overview.

[4] In addition to the Vendler classification, there are other possible analyses. Kenny (1963) proposed a three-way analysis in which "performances" (including Vendlerian achievements and accomplishments) contrast with activities and states. Smith (1997) presented a five-way analysis, which adds the semelfactive category to the four Vendler categories. Semelfactives are "single-stage events with no result or outcome" (p. 29) and include predicates such as *blink* and *cough*. In his paper on L2 acquisition, Robison (1995) advocated for a 6-way division: states, punctual states, activities, punctual activities, accomplishments, and achievements.

[5] Andersen and Shirai's (1996) formulation of the AH is cited more often than Andersen's (2002).

[6] More recently, Fuchs and Werner (2018, p. 212) provided evidence from corpus analysis showing that, although rare, progressive with statives was not categorically absent: The use of progressive with statives ranged from occurring with .22% to 1.53% of the statives produced by six L1 groups, and only erroneously extended to the range between .05% and .38% with statives. The same learners did use progressive with activities, showing that the low use with statives was not related to nonuse of progressive more generally.

[7] This seems to be a counterexample to the default past tense hypothesis, discussed later in this section.

[8] Gonzalez and Quintana (2018) only reported Spanish preterite and imperfect (although they stated they coded for present perfect). They also reported that they only coded recognizable forms for Vendler category, which is another reduction of data. Predicates that can be identified—regardless of morphology—should be categorized, and their morphology reported as "other."

[9] There are three items of each of 4 types (3 x 4), multiplied by the number of learners (N=112). The denominator should be the same for all the categories (336). By category, the missing data were: atelic + marker for imperfect (44 missing + 292 reported=336 items); telic + marker for preterite (51 missing + 285 reported =336); telic + marker for imperfect (49 missing + 287 reported =336); atelic + marker for preterite (50 missing + 286 reported =336). (See Tables 6-9, pp. 32-34.)

[10] Sugaya and Shirai (2007) focused on the inflection of –*teiru* (their dependent variable), which is a *nonpast* imperfective marker. In this respect, it is odd to examine the effect of the AH, which originally was posited for the inflection of past, examining the use of nonpast morphology. In addition to nonpast -*teiru*, Japanese has the past –*ta* and -*teita*, the latter being a past but not the exact past equivalent of -*teiru*, since the meaning of -*teita* is the past perfect, or past habitual (Sugaya & Shirai, Table 3, p. 6; Table 5, p. 13). Sugaya and Shirai provided data on the use of -*teita*, but only as an alternative for the use of -*teiru*. See Shirai (2002a, b) for a review of the association of –*ta* with achievements.

[11] L1 nonprogressive learners did not distinguish between progressive and resultative *–teiru*; L1 progressive learners used progressive *–teiru* more than L1 nongrogressive; and higher level learners (irrespective of L1) had more problems with the resultative state meaning).

[12] A middle ground is provided by personalized narratives in which speakers take the part of one of the characters in a film or book and tell the story from the character's perspective (Bardovi-Harlig, 2000).

[13] Tasks that generate too few tokens are also problematic. Using a 26-second film clip (Rastelli & Vernice, 2013), a 1-minute narrative (Muñoz & Gilabert, 2011), or 3 tokens per aspectual category (Diaubalick & Guijarro-Fuentes, 2017) may be equally fraught for issues of general task design. Attempting to make each of the 3 tokens clear contexts for the intended preterite or imperfect, Diaubalick and Guijarro Fuentes (2017) introduced adverbials in each context, which makes the test about the influence of adverbials or adverbials *and* lexical aspect, but not about the influence of aspectual category alone.