Salience and proficiency as determinants in the acquisition of L2 morphology

Francesco Romano

University of Essex, Department of Language and Linguistics
<francescobryanromano@gmail.com>

In a meta-analysis Goldschneider & DeKeyser (2001) showed five specific salience-related factors explain considerable variance in accuracy of L2 English morpheme use. This raises the interesting possibility that the predictive power of salience extends to languages whose morphemes are more salient than their English counterparts. Spanish and French learners of English and English and Spanish learners of Italian were compared for accuracy in use of five verb morphemes on two timed written tests. Only two of five factors, both top down salience factors, proved to be significant predictors of a multiple regression in the two L2s: perceptual salience and frequency. When additional bottom up factors related to salience, namely L1-L2 mismatch, L1 background, and L2 proficiency, were included, the latter was also found to be a good predictor cross-linguistically.[*]

KEY WORDS: salience, L2 acquisition, meta-analysis, morphology

1. Introduction

It is well-known in the language acquisition literature that second language (L2) learners omit or misuse some kinds of grammatical forms in production, despite evidence from input that such forms are obligatory for native speakers (e.g. bare verb forms in third person singular present or past tense contexts). One specific grammatical domain in which production does not correspond to the evidence available to learners is morphology. The morpheme order studies (Bailey et al. 1974; Dulay & Burt 1974; Larsen-Freeman 1975; Pica 1983; Rosado 1986), for example, showed that some grammatical morphemes are used accurately early on while others are not. Inaccurate use can persist into very advanced stages, even after considerable

[*] The research project reported was conducted under the auspices of the Italian Ministry of Foreign Affairs. I would like to thank Robert DeKeyser for his insights on this paper and four anonymous reviewers for constructive advice on prior versions of this paper. I am also grateful to the editors of the Italian Journal of Linguistics for their patience and suggestions. Any errors are mine.

exposure to naturalistic and fine-tuned input, leading to incomplete ultimate attainment (Lardiere 2006, 2007).

The task that has preoccupied researchers for many years has been to identify the specific factors which impinge on inaccurate use in production across levels of development. It is generally agreed that difficulty in the acquisition of L2 morphology is due to a combination of factors (Ellis 2002, 2006; Ellis & Larsen-Freeman 2006; Gass & Mackey 2002; Kwon 2005), and more notably properties of the target input (Goldschneider & DeKeyser 2001). Goldschneider & DeKeyser (2001) (G&D hereafter) have shown – with a degree of success – that five input factors all related to the concept of SALIENCE can explain substantial amounts (70%) of variance in accuracy scores obtained from twelve L2 English morpheme order studies of oral production compiled into a meta-analysis. The five salience-related input factors were: frequency, perceptual or phonological salience, morphophonological regularity, semantic complexity and syntactic category.

Salience as a construct was originally interpreted by G&D to refer to "the property of a structure that is perceptually distinct from its environment" (Ravin in G&D 2001). Ever since, new theories of salience have emerged, prevalently from the cognitive psychology field (Awh et al. 2012; Chiarcos et al. 2011; Docherty & Foulkes 2014). In this paper we re-operationalise salience by taking into account most recent theoretical developments in order to provide a more plausible implementation of the construct.

Given G&D's success in the identification of second language acquisition (SLA) determinants as properties all related to salience in the broader sense, a few important questions should be asked concerning the predictive power of the interaction and nature of such factors. First, it would be legitimate to ask whether the power of salience reflects in written tasks as much as oral tasks. If it does not, results would suggest, among other possibilities, the hypothesis that G&D's findings are limited to a 'performance' effect in oral production and may not be indicative of linguistic 'competence'. Furthermore, one might question whether the findings from G&D's (2001) meta-analysis of diverse morphemes would still apply if a more homogenous morpheme set were to be considered. Next, given salience factors are truly strong predictors of SLA in the morphological domain, then it would seem promising to explore the predictive validity of such a theory cross-linguistically (e.g. by investigating L2s whose morphological forms differ considerably from English with regards to salience). Finally, in light of the degree of success in predicting accuracy scores achieved by the five factors, it is worthwhile investigating other fac-

tors capable of explaining the remaining variance. Results of a statistical regression obtained from an L2 English study are compared to one obtained under methodologically equivalent conditions from an L2 Italian counterpart.

2. Determinants in the acquisition of morphology

That different properties of the input conspire to explain the acquisition of L2 morphology is well accepted in the literature. Ellis & Larsen-Freeman (2006), for example, have discussed the idea that SLA more generally (and morpheme orders more specifically) reflects complex, multivariate, non-linear interactions between variables. In best-case scenarios, any linguistic or non-linguistic variable may account for no more than 40% of variance in morpheme accuracy scores.

Ellis (2006) argues, for instance, that persistent difficulty in the acquisition of L2 morphemes is the result of the interaction of six factors, contingency, frequency, salience, interference, overshadowing and blocking, and perceptual learning, all shaped by the L1. In particular factors such as frequency, contingency, and salience, all tied to general cognitive processes of human and animal learning are claimed to be strong explanatory factors for L1 as well as L2 morpheme orders. According to Ellis, learning should be viewed in terms of a cue-outcome association whereby both the 'salience' of the cue and the importance of the outcome play a vital role in human or animal response to stimuli in the environment. To place this into perspective, grammatical morphemes are cues of very low (perceptual) salience:

> clitics, accent-less words, particles that depend accentually on an adjacent accented word and form a prosodic unit together with it, are the extreme examples of this: the /s/ of *he's*, /l/ of *I'll* and /v/ of *I have* can never be pronounced in isolation. Thus grammatical function words and bound inflections tend to be short […] with the result that these cues are difficult to perceive. (p. 170)

Factors such as L1 interference, L1 overshadowing and blocking, and perceptual learning, on the other hand, have an effect on how successfully L2 learners can tune their attention to form-meaning mappings in the L1 input.

The role of the learner's L1 as an explanatory factor has been also considered by Kwon (2005) and Luk & Shirai (2009). The authors

reviewed a set of morpheme order studies with the aim of exploring the predictive power of a learner's language background and grammatical transfer. Luk & Shirai (2009), for example, compared morpheme orders only in those studies where the learners' L1 was Spanish, Korean, Japanese, or Chinese, finding that differences between L1 and L2 in terms of the abstract category for morphemes like possessive *-s*, articles and 3SG *-s*, neatly accounted for differences in accuracy of use.

Gass & Mackey (2002) observed other factors related to 'innateness' may be implicated since frequency, considered for many years one of the best predictors of morpheme orders (Larsen-Freeman 1976; Ellis 2002), cannot alone explain linguistic facts (i.e. infrequent morphemes being acquired early on and with relative ease). The observation that "acquisition appears to proceed along its own route regardless of the input" (2002: 249), according to an innate blueprint, is a factor considered preponderantly in the morpheme order studies and a main tenet of some scholars adopting a generative approach to language acquisition.

3. The role of salience

G&D (2001) found that salience of input was crucial in explaining the morpheme order studies, downplaying the potential role of any innate predispositions to acquire language.

They focused on the predictive power of input properties in the acquisition of six grammatical morphemes, present progressive *-ing*, plural *-s*, possessive *-s*, articles, past *-ed*, and 3SG *-s*, by meta-analysing 12 L2 English studies. The studies selected elicited accuracy scores in use of the six morphemes via oral production from participants of various L1s, age, proficiency level, and learning the L2 in any of naturalistic, instructed and mixed mode settings. Therefore, a total of 72 cases, each representing the average accuracy score for an individual morpheme in any of the 12 studies, were obtained. In their meta-analysis, a linear multiple regression analysis comprising five salience factors (the independent variable predictors), explained a substantial portion of the variance in accuracy scores (the dependent variable) for the six morphemes: $R^2 = .71$. The R^2 value indicates that 71% of such variance could be accounted for by combining the values of five salience-related factors.

There are different theoretical and methodological aspects of G&D's study deserving attention. Before considering each of them

in turn, let us reconsider the notion of salience in light of recent research. Literature in the field of cognitive psychology suggests that (perceptual) salience is not only input or signal-dependent and should not be defined on the basis of surface features alone (Awh et al. 2012; Chiarcos et al. 2011; Docherty & Foulkes 2014). Salience, which is strictly dependent upon attention and memory and reflects "the degree of relative prominence of a unit of information at a specific point in time in comparison to other units of information" (Chiarcos et al. 2011: 2), is conceptualised in more than one way. Top-down approaches, for instance, stipulate that factors internal to the observer/speaker, which are independent from characteristics of stimulus proper, ultimately guide attention (Awh et al. 2012). In practice, when an individual searches for a friend's coat in a large room, it is the internally-motivated goal to locate the coat that determines its salience in the environment. Top-down approaches, nevertheless, fail to predict the fact attention can be disrupted by stimuli which are predominantly signal-dependent, such as a loud sound or sudden black out in the room, which make other cues more salient than the coat. These stimuli exemplify cases which make other properties of the input more salient (i.e. bottom-up approach). To contextualise for language learning, some examples of top-down properties of salience are factors such as: the L1, whereby the acquisition of a certain L2 morpheme may be influenced by several L1 morphophonological or orthographic constraints at a time; L1-L2 contrasts specific to the grammatical form being acquired (e.g. the L1 has no prefixing but the L2 does, the L1 allows null subjects while the L2 does not); L2 grammatical proficiency, whereby the internal grammar itself determines which forms are more or less salient or perhaps even accessible at different levels of proficiency. These factors contrast with signal-dependent bottom-up factors such as the input-related factors adopted by G&D. Ideally, salience should be operationalised taking into account as many top-down and bottom-up salience-related variables as possible, which is what we propose to do in this study.

3.1. Bottom up factors

The five bottom up salience-related predictors chosen by G&D were carefully selected among a range of candidates. Perceptual or phonological salience, the strongest predictor of the study, represents the easiness with which morphemes are perceived by the auditory system. By far the most discussed factor in the literature, a number of definitions for this term have been given, linked to both L1 (Brown

1973; Pye, 1980; Slobin 1971) and L2 acquisition (Dulay & Burt 1978; Ellis, 2006; Ellis & Larsen-Freeman 2006; Gass & Selinker 1994; Larsen-Freeman 1976). Brown (1973), for instance, claims children cannot learn what is not audible or of low perceptual salience, while Pye (1980) found that perceptual salience defined as "susceptibility to word and sentence stress and lack of disjuncture caused by a syllable boundary" (in G&D 2001: 23) predicts children's acquisition of person markers in Quiche Mayan well. All in all, the definitions considered agree that the easier a form is perceived in the input, the earlier it is acquired.

G&D operationalised phonological salience as the combination of three factors: number of phones, presence versus absence of a vowel (syllabicity), and relative sonority. In a recent study on the role of perceptual salience in the processing of Italian derivational morphology, Giraudo & Dal Maso (2015) operationalise perceptual salience as stress and size. Stress refers to whether or not a morpheme bears main word-level stress, whereas size amounts to the number of phones that make up the morpheme. It follows that stressed syllables within a morpheme and higher number of phones are associated to comparably higher salience.

The second factor, syllabicity, represents the presence or absence of one or more vowels in the morpheme, assuming presence increases likelihood of perceiving a form (G&D 2001: 23). Last is sonority which was originally equated to a sum of the scores for each phoneme as tallied by Laver's (1994) scale of sonority. However, a more fine-grained scale which also takes voicing of obstruents into account is the one proposed by Parker (2002) in Table 1. Although the scale is purported to be applicable to all languages, Parker validated his scale on experiments based of American English and Spanish, maintaining that, for experimental purposes, sonority scores can be quantified "as whole integers with identical distances between successive pairs" (Parker 2002: 273).[1]

Two factors were excluded from the G&D's computation, stress and serial position, since the English morphemes analysed do not bear main stress and were always at the end of the word with the exception of articles. Arguments for the exclusion of serial position as a component of phonological salience is given in Jany et al. (2007) which demonstrates that prosodic position of a phoneme within a syllable (e.g. onset or coda), located neither word initially nor word finally, does not alter its sonority in languages with very different consonant inventories.

Table 1. Parker's universal sonority hierarchy.

Phoneme	Score	Phoneme	Score	Phoneme	Score
h	18	r	12	k	6
vowels	17	w	11	g	5
y	16	l	10	ǰ, θ	4
m	15	š, f	9	n	3
s	14	š, f	8	z	2
č	13	p, v	7	t, b	1

The number of meanings expressed by a morpheme is its semantic complexity. Although there has been debate as to how to count semantic 'meanings', most L2 research testing semantic complexity as a determinant in SLA abides by Brown's hierarchy (Anderson 1978; Gass & Selinker 1994; Krashen et al. 1975; Larsen-Freeman 1976; Pye 1980; G&D 2011, among others). Brown (1973) established a hierarchy of complexity implying forms with more meanings are more complex, less salient than simpler ones and thus harder to acquire.

Affixation (free versus bound) and lexical/functional status determine a morpheme's syntactic category. Zobl & Liceras (1994) attempted to explain differences between order of morpheme acquisition in L1 and L2 acquisition by resorting to functional category theory. On this basis they proposed that free forms (e.g. *is*) are universally acquired earlier than bound ones (e.g. *-ed*), while lexical morphemes (e.g. uninflected words) are acquired earlier than functional ones (e.g. related to categories like D, T, C, I). In G&D's terms, bound forms are more salient than functional while lexical are more salient than functional (2001: 29).

The variable morphophonological regularity takes into account the degree to which phonological contours affect a morpheme, where those less affected are more 'regular', salient, and thus acquired earlier. Allomorphy, for example, has been shown to negatively enhance affixal salience in the processing of Finnish derivational suffixes (Järvikivi et al. 2006). In particular, lack of allomorphy, which amounts to invariance in the morphological structure of a word, increases affixal salience to the point of facilitating morphological decomposition. Although allomorphy, contractibility, and to some extent redundancy, are the factors most often cited in the literature, G&D (2001: 26) only considered the first sub-factor together with homophony, namely whether a morpheme is homophonous to any other functional forms. Redundancy, which characterises the repetition of meaning already conveyed by context (e.g. plural *-s* in *three apples*) was disregarded as it could not be controlled for in the meta-analysis.

Lastly, frequency represents a score for the token frequency of a morpheme in the input and was originally determined by calculating the token frequencies from a mother's speech-transcript in Brown (1973).[2] By far one of the most discussed predictors in the L2 literature (Brown 1973; Ellis 2006; G&D 2001: 29; Larsen-Freeman 1976; Larsen-Freeman & Long 1991, among others), the higher the frequency of a morpheme, the higher its salience and the earlier it is acquired. Naturally, in morphologically rich languages such as Romance languages, frequency of affix is almost always coterminous with the notion of productivity of verb class, whereby suffixes of the most productive class (e.g. Italian verbs ending in *-are*) are also the most frequent in the input.

3.2. Top down factors

Another noteworthy theoretical aspect of G&D's study concerns other top-down variables possibly implicated in explaining the remaining variance. In this respect, G&D (2001: 38) as well as other studies (Kwon 2005; Luk & Shirai 2009), suggest language background – the L1 system as whole – and L1 transfer – the possibility of one structure from the L1 affecting the acquisition of a near-equivalent structure in the L2 – can also act as significant predictor of SLA. Although the meta-analysis neutralised the effects of such variables by pooling results from various L1s, it was inevitably vulnerable to bleaching out the contribution these make to explaining variance.

Another relevant variable, for example, is L2 proficiency. The morpheme order studies considered cross-sectional samples only, measuring accuracy at one point in time (e.g. at a given proficiency level), but tended to generalise results to the whole population as long as scores for each morpheme 'crossed the finish line' (DeKeyser p.c.). A lack of longitudinal studies implies measurements at previous or subsequent stages of development may lead to ranks in accuracy scores different from the ones established, calling for studies where level of L2 proficiency is controlled for (Cook 1993; Ibanez 2012).

In summary, the three variables, language background, L1-L2 contrast for a grammatical property, and L2 proficiency, can be construed as salience-related endogenous top-down factors, all potentially implicated in explaining additional variance in accuracy of use.

3.3. Theoretical and Methodological issues

The finding that properties of salience were powerful predictors of the meta-analysis was particularly noteworthy for theoretical and methodological reasons. Theoretically, given that in SLA stud-

ies, factors taken individually are known to produce an R^2 which rarely exceeds .36 – that is they explain no more than 36% of a given linguistic phenomenon (Ellis & Larsen-Freeman 2006: 559) – G&D showed that significant amounts of variance could be explained by five salience factors. In turn, the explanatory power of such factors meant "no appeal to any blueprints or specific syntactic models is required to explain orders of acquisition" (G&D 2001: 36).

From a methodological standpoint, meta-analyses seek to establish the most significant statistical predictors for a given linguistic phenomenon from a large set of known studies which share the quality of answering identical or similar research questions. Subsequent to the publication of G&D's meta-analysis, though, no one has attempted to investigate the extent to which results of meta-analyses extend to single studies of L2s whose morphemes differ from English in terms of salience or even under different test modalities (e.g. written production) for example.

Langman & Bayley (2002) investigated the role of salience as defined by G&D in the acquisition of L2 Hungarian verbal morphology by nine adult L1 Chinese speakers. The authors were interested in the effects of the five salience factors identified by G&D on use of different tense and agreement morphemes in a morphologically rich L2, namely Hungarian. By completing a multivariate analysis of the learners' oral production, they found the probability a verb would be accurately inflected depended on the five properties, supporting G&D's claims that input properties tied to salience are strong determinants of the acquisition of L2 morphology.

Conversely, meta-analyses have limitations. One aspect of the G&D study – as well as the morpheme order studies more in general – that has come under direct criticism is the rather heterogeneous selection of morphemes investigated (see Ibanez 2012: 22 for a discussion). Investigating fewer closed-class categories (i.e. only V or only N inflectional forms) is, therefore, crucial to ascertain whether the predictive validity of salience was not purely a 'mechanical' effect resulting from the choice of morphemes.

A final consideration rests with cross-linguistic plausibility. If salience factors turn out to be good predictors for the acquisition of morphology in L2s beyond English, its theoretical importance would be significantly increased. In fact, on the basis of G&D's findings, salience should be strongly predictive in any type of language, regardless of its morphology being more or less salient than English, even though languages where this difference is more striking offer a prime test case. Selecting a comparable set of morphemes cross-linguistical-

ly, however, seems a good starting point to minimise different operationalisations.

4. Research Questions

In this study, we compare the acquisition of L2 English morphemes to L2 Italian near-equivalents, with the expectation the five input-related factors will be statistically significant predictors for both L2s. Our study extends G&D's findings by measuring the extent to which the five salience factors: (1) predict performance in L2 written production; (2) predict accuracy scores of a closed set of L2 morphemes; (3) are cross-linguistically plausible; (4) co-participate with previously unexplored variables in explaining variance. If the predictive validity of the salience factors reflects also in written production and is cross-linguistically plausible, by virtue of accounting for large amounts of variance despite conspicuous differences in salience properties between the morphology of the two target languages, important implications can be drawn.

5. Method

5.1. Participants

A total of 62 participants took part in the study. Of the 30 adult instructed learners who participated in the L2 English study, 14 were L1 French speakers and 16 were L1 Spanish speakers. Of the 32 participants to the L2 Italian segment, 17 spoke Spanish and 15 spoke English as L1. The age of participants in both studies ranged from 17 to 53. The French and the Spanish-speaking participants to the English study were living in the UK and attending English language courses at private or public language schools at the time of study. Likewise, the Spanish-Italian and English-Italian learners were living in Italy and attending Italian language courses at either private or public language schools. The effect differences in proficiency at time of study would have on the results, were estimated via a modified version of the online version of the 'structures' component of the DIALANG test (http://www.lancs.ac.uk/researchenterprise/dialang/about). This test was preferred to distinct English and Italian placement tests such as the Oxford Placement Test and samples of the Certificato Conoscenza Lingua Italiana (CELI) respectively because, at the time, DIALANG was probably the only diagnostic test designed

to relate L2 learners' scores for both languages directly to a level of the Common European Framework (CEFR) (Alderson & Huhta 2005). Due to the limited data being available on the validity and reliability of the Italian version (Alderson & Huhta, 2005: 310) and its perceived difficulty (Reigen 2011), the version available online was pretested as a paper-and-pencil test on a sample of 20 L2 speakers during a pilot test phase. Before administering the test, the researcher, a native speaker of the Italian, completed a face validity analysis of 30 items, identifying a few of them which required revision. The final version was composed of 30 items divided between gap-fill, multiple choice, and sentence completion techniques. Generally the participants performed well and did not find the new test too difficult in both the pilot and subsequent main phase. In Table 2 I report for each group mean L2 proficiency, as well as length of instruction and exposure to the L2 for purely informative purposes.

Table 2. Participant data.

L1	N	L2	L2 Proficiency Mean, (SD), and range	Mean L2 instruction (years)	Mean exposure (months)
Spanish	16	English	.71 (.15) .33-.90	8.5 (2.1)	21.83 (5.9)
French	14		.72 (.13) .47-.90	7.27 (1.7)	6.44 (1.3)
Spanish	17	Italian	.56 (.16) .25-.81	1.68 (.55)	3.87 (.9)
English	15		.52 (.21) .27-.80	1.91 (1.2)	6.81 (1.2)

Generally, longer periods of instruction in years corresponded to higher proficiency but not longer periods of exposure. Of utmost importance for the multiple regression analysis was that all groups showed a sufficiently wide range of scores, reflected in the fourth column by the range row. Five of the Spanish-Italian participants who also had knowledge of English also took the English proficiency test but their scores were low (mean .43, SD .03, range 0 to .50), with only one participant scoring .50.

Twelve adult English and Italian controls were recruited respectively in England and Italy and tested on both the L2 proficiency test (an L1 test in their case) and the main tests. Proficiency scores and main test scores were 100%.

5.2. Operationalisation

The accuracy scores for five morphemes, namely 3SG and 3PL uncontractible copula *is* and *are*, two 3SG *-s* allomorphs perceived as /s/ and /z/, and past participle perfective *-ed* perceived as /t/, all belonging to the verbal domain, constituted the dependent variable (DV) for the L2 English study. In order to test cross-linguistic plausibility of G&D's claims, the DV accuracy scores for the following five near-equivalent L2 Italian inflections were selected: 3SG and 3PL copula *è* and *sono*, 3SG productive *-a* for verbs of the first class ending in *-are* in their infinitive form and its allomorph *-e* for verbs of the second and third class ending in the infinitive form *-ere* and *-ire*, and the past participle perfective form *-ato*.

Selecting a comparable set of morphemes cross-linguistically is not simple. The two target languages, Italian and English, differ not only for salience, but also for other morphological reasons. For instance, choice of the appropriate phonological form for the 3SG morpheme *-s* in English is phonologically-conditioned (compare *sell*-/z/, *reflect*-/s/, *finish*-/ɪz/) and intrinsically dependent upon the phonological qualities of the stem's coda (affricate, labial, nasal, etc.), whereas in Italian the choice of a 3SG allomorph is lexically-conditioned (compare *am*/a/, *tem*/e/, *dorm*/e/ where the coda of the roots is equally nasal in all three cases). In addition, verb formation in English is almost always additive (*reflect* > *reflect-s*, *reflect* > *reflect-ed*, *reflect* > *reflect-ing*) while in Italian it may be either additive or subtractive depending on whether the native or L2 speaker takes the root, the stem, or the infinitive as the starting point for inflection (*am* > *am-a*, *am-a* > *am-a-to*, *am-are* < *am-a*).

Although the above cross-linguistic differences suggest a disparity in terms of the morphological processes involved, the choice of morphemes is dictated primarily by a need to establish the same if not better degrees of statistical variation in the scores for each morpheme compared to G&D's method. 3SG /s/, for example, differs from its allomorph /z/ in ways important for the operationalisation of salience and cross-linguistic plausibility. First, in relation to salience, the two forms differ with respect to sonority and frequency in the input. In terms of sonority, the two phonemes differ for voicing: /s/ is voiceless, /z/ is voiced. If the frequency of 3SG *-s* in the input is counted, the frequency when sounded as /s/ is clearly different from when sounded as /z/ (see Table 3). Crucially, such differences could be responsible for the well-attested omission phenomenon for this morpheme, whereby rate of omission for one allomorph may be higher than the other, despite the fact both spell out orthographically as *-s*.

This juxtaposition was also necessary to create symmetry with the Italian 3SG allomorphs *-a* and *-e*. At present, the L2 English and L2 Italian studies reported here is the first attempt in the literature to compare morphemes cross-linguistically, under methodologically equivalent conditions.

The operationalisation of bottom-up salience factors in this study was in many aspects identical to that in G&D's meta-analysis (2001: 25) with the exception of a few amendments. The phonological salience score was based on stress, in addition to number of phones, syllabicity, and sonority. Morphemes bearing main stress like *-ato* in Italian receive a score of 1 in contrast morphemes that do not like *-a* which receives 0. Number of phones represented the total number of phonemes per morpheme, while syllabicity distinguished absence (0) or presence (1) of one or more vowels per morpheme. Sonority was tallied following Parker (2002) and not Laver (1994), taking into account additional factors such as voicing, resulting in a more fine-grained and wider scale compared to the one utilised by G&D (recall Table 1). Thus, 3SG copula *is* scores 20, 17 for the vowel /I/ and 3 for the consonant /z/. As discussed above, semantic complexity was the sum of the meanings expressed by a morpheme, where more complex forms are less salient than simpler ones. On the other hand, affixation and lexical/functional status determine syntactic category for which score and salience are directly proportional. Scores were assigned in the following way: lexical free = 4, lexical bound = 3, functional free = 2, functional bound = 1. Morphophonological regularity was a composite score based on allomorphy plus homophony, the assumption being that their relationship to salience is inversely proportional: the higher the score, the less salient to the learner. Thus, as discussed previously, the lower the degree of affixal allomorphy, the more invariant the morphological structure of the word, and the higher facilitation of morphological decomposition (Järvikivi et al. 2006). Finally there is frequency which was established as in G&D (2001: 30), randomly selecting three sets of parent-child speech data transcribed for the CHILDES (MacWhinney 2000) corpus and conducting analysis via the CLAN tool. The English sets were taken from the Brown's (1973) corpus while this Italian sets were drawn from Romance-MOR/Romance for Italian. The English scripts were adam25, eve18, sarah018, whilst the Italian were gre4 (Tonelli corpus), cam211 (Antelmi corpus), and dia09 (Calambrone corpus). Although baby talk is not ideal for comparison to the speech directed to adult L2 speakers, the morphemes considered are generally so frequent in speech that minimal variation is expected to occur across data sets.

Randomised selection of the sets was especially intended to minimise such effects. The operationalisation of the dependent and independent variables is presented in Table 3.

Table 3. Salience of five Italian and English morphemes.

Morpheme IT vs. EN	Perceptual Salience				Semantic Complexity	Syntactic Category	Morphophonological Regularity		Frequency
	Phones	Sonority	Syllabicity	Stress[1]			Allomorphs	Homophony	
è vs. *is*	1/2	17/20	1/1	1/1	3/3	2/2	1/1	2/2	104/111
sono vs. *are*	4/1	52/29	1/1	1/1	3/3	2/2	1/1	2/2	12/8
3SG.PRES *-a* vs. /*s*/	1/1	17/14	1/0	.5/.5	3/3	1/1	2/3	2/2	51/35
3SG.PRES *-e* vs. /z/	1/1	17/3	1/0	0/.5	3/3	1/1	2/3	2/2	21/19
PAST.PART *-ato* vs. -/t/	3/1	36/2	1/0	1/0	4/2	1/1	3/2	2/2	6/5

Note. [1] 3SG *-a*, /s/, and /z/ receive half a point because they only bear stress as a suffix to one of the two stems tested for each, *va*, *checks*, and *goes* respectively.

The salience properties of English morphemes are visibly different from the Italian near-equivalents. The shaded cells highlight pairings where the Italian morpheme is more salient than the English counterpart, showing a much larger number of cases in which Italian forms surpass the English and not vice versa (15 versus 5). This is especially true of perceptual salience, for which the Italian forms are more salient than the English on all four sub-factors.[3]

As for the top down factors, the data from the ESL studies used in G&D (2001) originally included samples whose participants varied in age, L1, and mode of acquisition (i.e. naturalistic, instructed and mixed mode settings) and did not control for effects of L2 proficiency. In this study, L1-L2 differences for type of agreement affix, the learners' L1, and proficiency at time of study, all top down salience factors, were operationalised as the L1-L2 mismatch, 'language background' (L1), and L2 grammatical proficiency independent variables respectively.

The L1-L2 mismatch variable was necessary to account for a difference between the agreement systems (or null subject property) of the L1 and L2. Verb affixes in null-subject languages are more salient than their English and French counterparts for morphosyntactic reasons: they act as incorporated subject pronouns. Under some theo-

retical accounts (Bresnan 2001; Dalrymple 2001), agreement affixes on verbs in languages such as Italian, Spanish and Chichewa also encode predicate subject function as a consequence of pronominal subject omission in the appropriate pragmatic and syntactic contexts. Thus, the correlation in consistent null subject languages between possibility of subject omission in finite contexts and richness of morphology is not coincidental: it amounts, in formal linguistic terms, to the option of the subject surfacing as a verb suffix rather than a full pronoun in preverbal finite position as in English and French. A match or mismatch between L1-L2 with respect to pronominal incorporation was operationalised as an L1-L2 mismatch independent variable. Consequently, L1-L2 mismatch stipulates a score of 0 when there is a mismatch between the type of agreement in L1 and L2 but a 1 when there is a match.

Language background is simply a categorical variable used to classify learners as belonging to either a Spanish speaking (0) or French speaking group (1) for L2 English, and English speaking (0) or Spanish speaking (1) group for L2 Italian. Lastly, the L2 proficiency variable is an independent value measuring the learner's L2 grammatical proficiency at time of study by means of a placement test which was composed of 30 items testing L2 grammar and lexicon. Each participant's proficiency score, thus, ranged between 0 and 30.

5.3. Materials

One aim of this study was to test whether salience factors are equally good predictors of accuracy in use of morphology in target languages other than English. Establishing cross-linguistic comparability, both in terms of the morphemes to be tested and the conditions under which they are to be elicited is not an easy task. Spontaneous oral tests of the kind employed in the morpheme order studies, for example, are unlikely to elicit morphemes in identical sentences across L2s. The sentential contexts (subject, agreeing verb, and the relative syntactic distance from each other) in which morphemes are elicited need be carefully matched. Controlled elicited oral imitation tests or written tests, where all parts of the sentence can be translated more easily from one language to the other, overcome such limitation.

Other potential issues in designing tests for this study are related to participants avoiding use of the forms needed by the researcher, producing them an insufficient number of times, and only producing them for a limited set of lexical stems. The latter is especially important for a study of L2 morphology because L2 learners may be applying appropriate morphological rules only to certain stem types or may

be utilizing memorized chunks, erroneously giving the impression of acquisition. For an in-depth discussion of such limitations in oral production tests, for instance, see the productivity and systematicity analyses in Pienemann (1998) who suggests analysing different types of stems to which affixes attach in order to avoid erroneously drawing inferences from use of memorized chunks.

Next, use of a battery of tests is necessary to ascertain whether the predictive validity of salience is limited to oral production. It may in fact be that the findings of G&D are due to a performance, rather than a competence, effect of meta-analysing scores obtained from oral production only. If accuracy scores are a reliable indication that L2 learners have knowledge of the abstract agreement features mapped onto morphemes and can distinguish among the phonological/orthographic forms realising them, then triangulation across tests of different modalities is needed. Oral task effects due to pressure on access and retrieval of a lexical form for a given morpheme (e.g. /s/ /z/ /ɪz/ for 3SG) may yield lower scores, incorrectly suggesting that L2 learners: (1) have not acquired the abstract features and their associated phonological realisation (see the Missing Surface Inflection Hypothesis of Prévost & White, 2000, for this view); (2) salience is a good or bad predictor. Thus, in order to reliably gauge whether salience is a good predictor of variance in accuracy scores and knowledge of morphology, results employing other methods need be compared to oral production (see the morpheme order study conducted by Larsen-Freeman, 1976, for a similar implementation).

One disadvantage of written tests compared to oral production, however, is the former do not permit direct elicitation and analysis of each of the phonological forms associated with 3SG *-s*. It could be argued, for example, that the three allomorphic variants of 3SG *-s* cannot be tested via written production because two of them, /s/ /z/, map onto the same orthographic form, the letter <s>, while /ɪz/ maps onto *-es*. In contrast, oral production allows elicitation of each of the allomorphic variants if the audio recording tools utilised and the transcriber employed are sufficiently precise.[4]

Two speeded written tests, a gap-fill and a sentence completion, were employed. The tests were designed so as to accommodate a wide range of proficiency levels.[5] For this reason, the main tests used simple vocabulary and were made available in two versions, one with instructions written in the learner's L1 and one in the L2. All participants, nevertheless, chose to complete the L2 version. A picture of a busy city main street taken from Ashworth & Clark (2005) depicting various scenes was presented to participants. The test items, intro-

ducing either a one-word gap in the gap-fill test or a short phrase gap in the sentence completion, described real-life situations occurring in the picture in the form of narrative-descriptive text.

In the gap-fill there were in total 24 items subdivided in 10 test items and 14 fillers. The 10 test items comprised 2 items for each of *is*, *are*, 3SG /s/, 3SG /z/, and past participle *-ed* English morphemes and 2 for *è*, *sono*, 3SG *-a*, 3SG *-e*, and *-ato* in the Italian version. To test each of these morphemes, participants were asked to choose a form, inflected or uninflected for one of the following bare verb forms supplied in parenthesis: *essere* (copula forms), *andare* and *controllare* (for 3SG *-a*), *vendere* and *riflettere* (for 3SG *-e*), *essere andare* (for *-ato*); *be* (copula forms), *go* and *sell* (for 3SG /z/), *check* and *reflect* (for 3SG /s/), *have finish* (for *-ed*). Table 4 classifies each of the verb forms for their token frequency, transitivity, and allomorph (English) or verb class (Italian).

Table 4. Predicate candidacy.

Language	Stem or Word Form	Token Frequency	Transitivity	Class or Allomorph
English	*is/are*	high	-	/
	go/goes	high	-	z
	check/checks	high	+	s
	sell/sells	medium	+	z
	finished	high	+	/
	reflect/reflects	low	-	s
Italian	*è/sono*	high	-	/
	va/vanno	high	-	-are
	controlla/controllano	high	+	-are
	vende/vendono	medium	+	-ere
	finito	high	+	/
	riflette/riflettono	low	-	-ere

Note: + = transitive, - = intransitive. / = not of interest.

Token frequency was obtained via the lexical dictionary of De Mauro et al. (1993), according to which forms *va* and *vanno* of *andare* have the highest token frequency, followed by *controlla* / *controllano* of *controllare*, *vende* / *vendono* of *vendere*, and *riflette* / *riflettono* of *riflettere*. These frequencies are consistent with verbs of the first class ending in *-are* being the most productive. Following the Corpus of Contemporary American English, the English lexical verb *check* / *checks* has the highest token frequency, followed by *go* / *goes*, *sell* /

sells, and *reflect / reflects* (Corpus of Contemporary American English). Thus, unlike the relationship between frequency and productivity of verb class in Italian, there appears to be no correlation between token frequency and allomorph type in English (i.e. stems requiring /s/ are no more frequent than those for /z/). Finally, there were two verbs for each allomorph type in English or verb class in Italian, one transitive and one intransitive.

Examples (1a) and (1b) illustrate elicitation of copula *essere* 'to be' in Italian and English respectively, while elicitation of the lexical verb *riflettere* 'to reflect' is illustrated in (2a) for English and (2b) for Italian.[6]

(1) a. *Nella cabina telefonica l'uomo _____ (essere) al telefono.*
b. *In the phone booth, the man _____ (be) on the phone.*

(2) a. *Il commerciante _____ (riflettere) sulla giornata lavorativa che lo attende.*
b. *Every day the shop-keeper_____ (reflect) upon the busy day ahead of him.*

One item testing (*is*, *are*, and 3SG /s/ in the L2 English version, *è*, *sono*, and 3SG *-a* in the L2 Italian) also required a decision on word order between an adverb of frequency and the verb to further distract participants from inflection. I report an example for copula *essere* 'to be' in (3):

(3) a. *I trasporti in città sono efficienti ma nel weekend molte persone si lamentano.*
_____ sempre (essere)_____ scontente il Sabato e la Domenica.

b. *Transportation in the city is efficient on weekends. Many people complain because there are not enough buses and trains.*
They _____ usually (be)_____ unhappy on Saturdays and Sundays.

Shorter and simpler items such as those testing copula inflection by itself (example (1)) were placed at the beginning, whereas longer and more complex items such as those testing morphemes in conjunction with word order (example (3)) were placed towards the end.

The 14 fillers were 4 items testing vocabulary, 2 relative pronouns, 6 3PL lexical verb forms, and 2 present progressive forms (replaced by 2 present perfect forms in Italian with the main verb ending in *-ito*). Vocabulary fillers came before all other items in order

to lead learners to believe the test would require attending to word meaning rather than word form. In contrast, fillers testing relative pronouns were the very last two items to be completed as these were the most complex. Fillers eliciting knowledge of 3PL morphemes were inserted throughout. Differentiating fillers had the effect of leading participants to believe the test was examining a variety of structures. An example for each filler type is given in examples (4) through (7) respectively:

(4) a. *Vicino al gatto, davanti ai veicoli, c'è un giovane ragazzo che indossa una maglia a righe. Usa una_____ perché è stanco di guidare la macchina nel traffico.* (L2 Italian, vocabulary)

b. *Next to the cat, ahead of the vehicles is a young man wearing a striped jumper. He is riding a_____ to work today because he is tired of driving his car in traffic.* (L2 English, vocabulary)

(5) a. *Ci sono due negozi di giocattoli su questa strada. Il negozio di giocattoli _____ apre presto è pieno di bambini.* (L2 Italian, relative pronoun)

b. *There are two toy-shops on the high Street. The toy-shop _____ opens early is filled with children.* (L2 English, relative pronoun)

(6) a. *La Domenica, generalmente due signore _____ (andare) a comprare la frutta* (L2 Italian, 3PL)

b. *Every Thursday two ladies _____ (go) there to buy some fruit.* (L2 Italian, 3PL)

(7) a. *C'è un uomo che sembra molto stanco. Lui_____(finire) da poco di fare una lunga camminata.* (L2 Italian, Present Perfect *-ito*)

b. *There is a man who looks very tired. He_____just_____ (finish) walking across the city.* (L2 English, Present Progressive)

In both the English and Italian tests, one item testing 3PL inflection (a bare verb form in English) also tested knowledge of word order between an adverb of frequency and the item verb to further distract participants. The English version of the test was exactly one minute longer than the Italian (8 versus 7 minutes).[7]

The same sentences embedding the test items, and test items

but slightly different fillers were used in the second test, the sentence completion. By maintaining the design of the items consistent across tests, the effects of differences in presentation of the items were minimised. For instance, if the type of verb and agreeing subject were varied across tests to examine 3SG *-a* in L2 Italian, *amare* in one test but *dominare* in the other, variance in accuracy rates might have been explained by item effects. In this way, any substantial differences in the results could more reliably be attributed to either between-participant or between-test effects.

The mechanics of the sentence completion test differed from the gap-fill in that participants were not told which subject and agreeing verb to use. Instead, a bare verb form had to be first chosen from an item bank containing the necessary verb plus a series of distracter words. Next, the participant had to decide whether to inflect the verb and add a subject. Six of the Italian test items and all 10 of the English required a subject. In this way the degree of control over participant responses was significantly decreased compared to the other test. Examples (8) and (9) of the sentence completion correspond to the previously illustrated examples (1) and (2) taken from the gap-fill:

(8) a. *Nella figura c'è un taxi e un uomo in una cabina telefonica. Il taxi sta fermo.* _____ *è al telefono.*

b. *In the phone booth,* _____ *on the phone.*

(9) a. *Vicino al lampione c'è un uomo con i baffi e un commerciante. L'uomo con i baffi pensa a sua moglie.*_____ __ *sulla giornata lavorativa che lo attende.*

b. *Near the lamp-post there is a man with a briefcase and one with a moustache. Every day* _____*upon the busy day ahead of him.*

In completing sentence (8b), for example, the participant noticed a man in a phone booth when looking at the picture. The item bank provided, among other options, the uninflected verb *be* to which a suiTable subject needed to be added (e.g. *a man*, *the man*, *he*).

The purpose of requiring a subject to be added in null subject Italian was two-fold. First, it enables a fairer comparison to the English equivalent sentences which ban null subjects. If the contexts were not controlled for, i.e. allowing both null and overt subjects in Italian, the distribution of subjects would not be comparable across the L2s due to differences in discourse contexts (null subjects are only

possible with topic continuity). Second, requiring realisation of the subject allows one to tease apart errors of anaphora resolution from correct use of inflection. L1 Spanish learners of English, for example, may produce incorrect null subjects but appropriate verb inflection due to L1 transfer. By the same token, any group of participants may produce a subject misagreeing for number, person, gender features with its referent antecedent (e.g. *they* in a 3SG context) yet correct inflection on the verb in the given context (e.g. *reflects* in 3SG) (see also Poeppel & Wexler 1993 for this methodology).

Given the increased complexity of the task, the vocabulary fillers were eliminated, bringing the total number of items down to 20, 10 of which were test items and 10 fillers. To further simplify the test, it was divided in two parts: part 1 with 6 items and part 2 with 14 items. If the 20 items were presented at all once, the respondent would have had 24 options to select from for the first item, 23 for the second, and so on.

Part 1 comprised 6 items composed of 3 test items and 3 fillers. The test items were a 3SG copula, a 3PL copula, and a 3SG verb form, all tested in conjunction with word order, while the three fillers were the 2 relative pronouns plus one 3PL item. In addition to the 6 responses necessary to complete the items, there were an extra 4 distracters, 2 involving ungrammatical word order choices with copula *essere* 'to be' and 2 with lexical verb *controllare* 'to check' (e.g. *sempre essere* and *sempre controllare* where the order of the adverb is incorrect with respect to the verb). The L2 Italian and English versions were identical in all respects, including the timing (6 minutes).

Part 2 comprised 14 items, equally divided in test items and fillers. The 7 test items were broken down as follows: copula 3SG and 3PL forms (1 each), the 3SG /s/ and *-a* (1 item), the 3SG /z/ and *-e* (2 items), and the past participle *-ed* and *-ato* (2 items). The fillers tested use of 3PL inflection (5 items) and the present progressive (2 items). In addition, 2 distracter words, *read* and *realise*, unnecessary for completion of the items were also provided in the item bank to deter the respondent from focussing purely on word form. The Italian test Part 2 was identical in all respects to the English except the timing (8 minutes in English versus 11 in Italian).

The order of presentation of the items was the same for both tests. Randomizing the item order of presentation would not have enabled participation of lower proficiency learners, who under the current design, were gradually eased into more difficult items. A facility value analysis of test items placed at the beginning and end of the test confirmed no bias for order of presentation insofar as the average

score on items testing a copula and a 3SG morpheme at the beginning and at the end of the test was always nearly identical. This held true for both tests, L1 groups, and L2s.

5.4. Procedure

Participants met the investigator either individually or in groups on school premises. Every participant completed in the following order, an ethical consent form, a personal data questionnaire, the proficiency, the gap-fill, and the speeded sentence completion test. The three tests were paper-and-pencil and were completed under the investigator's supervision whose main responsibilities were to clarify instructions if necessary and stop participants once the time for completing each test had run out. Participation in the study was voluntary. Five practice blocks were worked through and questions answered before commencement of the experiment.

6. Results

6.1. L2 English

According to G&D (2001), the five salience factors are expected to play a significant role in explaining variance in accuracy of use of a closed set morphemes. Moreover, if the role of salience were also substantiated in written tests, this would suggest qualities of the input have deeper consequences for the representation of morphological knowledge in L2 learners. The accuracy scores for the five morphemes were calculated as suppliance in obligatory contexts (SOC) in the two written production tests. The average participant score from one test was then averaged with the other, obtaining a final overall average for each morpheme reported in Table 5. SOC scoring was common to all the studies meta-analysed by G&D and considers three types of answers: suppliance (scoring a 1), no suppliance (scoring a 0), and incorrect suppliance (scoring .5). To exemplify, the perfective morpheme *-ed* scored a full point if supplied in obligatory contexts (OCs) while incorrect inflection (i.e. *-ing* in an *-ed* context) scored .5. Likewise, suppliance of the *-s* form scored a 1 if in 3SG contexts, while *-ing* in the same context scored .5. Following Poeppel & Wexler (1993), the scores for the 3SG morpheme in its two allomorphic variations scored 1 even if the subject was incorrectly omitted or mismatched in agreement features, provided the verb stem and inflection were correctly spelled. In a 3SG OC an answer that took the form of *Ø reflects*, where the subject is incorrectly omitted or *they reflects*, where the

subject is supplied but inconsistent anaphorically with the context, scored a 1 in either case because the respondent provided the right verb form. Responses like *he reflectes* or *Ø reflectes*, instead, scored .5 because the affix is the incorrect allomorph.[8] There were in total 10 OCs in each test, 2 for each *is*, *are*, /s/, /z/ and *-ed*.

The final average accuracy scores for the Spanish and the French participants are presented separately in Table 5.

Table 5. L2 English average accuracy scores.

L1	Rank				
	1	2	3	4	5
Spanish	*are* .89/.20	*is* .70/.24	*-ed* .67/.33	/z/ .56/.39	/s/ .35/.36
French	*is* .82/.15	*are* .80/.26	/z/ .71/.30	*-ed* .60/.34	/s/ .41/.33

Note. Values represent mean and SDs, from left to right of the bracket. Ranks do not differ by test.

The mean accuracy scores reported under each morpheme are ranked from highest to lowest from left to right. Ranks differ somewhat by L1 in light of the position of the free morphemes *is* and *are* scores being inverted and *-ed* ranking lower in the French group. However, if scores for the SG and PL copula forms are averaged to obtain a single copula score - as was done in the morpheme order studies, for instance - the difference in rank no longer subsists (L1 Spanish mean copula score = .80, L1 French mean copula score = .81). Both groups found 3SG *-s* to be the most difficult when the corresponding phonological form to be perceived in the input is the /s/ allomorph. This is a new finding in the context of studies investigating L2 English morphology given that no studies in the past, to the best of my knowledge, have ever differentiated accuracy scores on 3SG *-s* based on the corresponding allomorph in the input. Table 6 reports the L2 English scores by L1 group for each of the 4 predicates averaged across the two tests. Qualities considered for predicate selection, previously described in Table 5, are repeated here.

Table 6 suggests neither token frequency nor transitivity can explain the rank for average accuracy levels. If high token frequency of the stem were relevant then *go*/*goes* and *check*/*checks* would have occupied position 1 and 2 of the rank; instead both Spanish and French participants were fairly accurate in their use of *sell*/*sells* which has medium stem token frequency. Nor does the transitivity of a predicate imply higher ranking: intransitive verbs *go*/*goes* and

reflect / reflects are the most and least accurate type respectively in the ranking of each L1 group. It appears instead that the allomorph of the 3SG affix selected by the stem better predicts accuracy scores but without computing inferential statistics to confirm the effects described this conclusion remains tentative. Whether the descriptive statistics in Table 6 can be validated by inferential statistics is an issue we leave open for future research since it is beyond the scope the study to investigate linguistic factors beyond the five salience factors proposed by G&D and the three new variables proposed in the literature considered in the study.

Table 6. L2 English average accuracy scores by predicate type and their qualities.

Stem or Word Form	Token Frequency	Transitivity	Allomorph	Average Score/ Rank	
				Spanish	French
go/goes	high	-	z	.57/1	.64/2
sell/sells	medium	+	z	.53/2	.78/1
check/checks	high	+	s	.40/3	.49/3
reflect/reflects	low	-	s	.30/4	.33/4

Note. + = transitive, - = intransitive.

Overall, the ranks found are consistent with those in Pica's (1983) study of L1 Spanish learners of English & Krashen et al.'s (1976) study of French learners of English, if the choice of past regular *-ed* in those studies and past participle *-ed* in this study is set aside: *is* (1), *-ed* (2), *-s* (3). Finding the same orders as Pica & Krashen et al.'s studies (both based on oral data and meta-analysed by G&D) speaks to the similarity of results using the written and oral modes of elicitation.

Two multiple regressions were computed, one with the five top down salience factors alone, the other with the three bottom up factors added. Table 7 summarises the results of these analyses.

The table's top portion details results for a pure replication of the model made up of the 5 top down salience factors in G&D (Model 1) while the bottom reflects results for the model to which 3 new bottom up non-linguistic variables were added (Model 2). The Sig column, indicating the significance test results for each variable, reports the degree of reliability or meaningfulness for each regressor in terms of explaining variance in the DV. All of the top down factors except syntactic category were significant to both models, but in model 2 L2 pro-

ficiency and L1 background are also good predictors. The values for the five salience factors in model 2 are not reported for conciseness but are only slightly different from model 1. The ß column defines the unique contribution of each IV to the regression model once scores are standardised. In model 1, for example, morphophonological regularity has the greatest impact on the DV with a value of -.83. By themselves, the five variables explain close to 17% of the total variance (expressed by the R^2 value) in Model 1, but when the new variables are added in, the amount of variance explained improves to 39%. Four of the 5 salience factors are found to significantly predict variance in the study of L2 English.

Table 7. Multiple regression L2 English.

Model	#	Variable	B	SE B	ß	t	Sig
	1	Morphonological regularity	-.72	.15	-.83	-4.68	< .001
	3	Semantic complexity	.53	.14	.69	3.71	< .001
1	2	Perceptual Salience	-.08	.02	-.64	-2.90	< .01
	4	Frequency	-.00	.00	-.26	-2.78	< .01
	5	L2 Proficiency	.01	.00	.43	7.35	< .001
2	6	L1	.19	.07	.27	2.64	< .01
	7	L1-L2 Mismatch	.13	.09	.18	1.51	.13

Note. SE = standard error. Method: forced-entry in SPSS, n = 150. For model 1, R = .41. R^2 = .17. For model 2, R = .62. R^2 = .39. Syntactic category was excluded by SPSS 19 for failing collinearity requirement.

6.2. L2 Italian

The second prediction that salience factors would also play a significant statistical role in explaining variance in use of L2 Italian functional morphemes was tested. Accuracy scores for the five morphemes copula *è* 'is', *sono* 'are', 3SG *-a*, 3SG *-e*, and the past participle *-ato* were calculated adopting the same methodology described above for the L2 English morphemes. Given omission phenomena are not possible in Italian, we considered infinitive forms comparable to English bare forms, scoring them with a 0 (see Prévost & White 2000 for the same implementation in L2 French). Other non-target-like suffixes scored .5 while target-like forms scored 1. The accuracy scores for the English and the Spanish participants are presented in Table 8.

Table 8. L2 Italian average accuracy scores.

L1	Rank				
	1	2	3	4	5
Spanish	*sono* .82/.26	*è* .80/.16	*-a* .61/.37	*-ato* .47/.42	*-e* .41/.26
English	*sono* .85/.15	*è* .73/.29	*-a* .66/.30	*-ato* .60/.35	*-e* .43/.29

Note. Values below the morphemes represent mean and SDs, from left to right of the bracket. Ranks do not differ by test.

The Spanish and English orders found are the same. Both groups scored highest on copula *sono*, followed by 3SG *-a* which attaches to the roots of the first productive class, forming stems very high in token frequency (recall Table 5), and ending with the 3SG *-e* which affixes roots of the second class, forming stems with lower token frequency. More interestingly, comparing the L2 English (Table 5) and L2 Italian (Table 8) ranks suggests there is cross-linguistic evidence that free morphological forms are used more accurately than bound ones.

Table 9 summarises the predictive power for the 5 top down regressors (model 1) as well as the power of the same regressors plus the three bottom up factors (model 2) predicting L2 Italian accuracy scores.

Table 9. Multiple regression L2 Italian.

Model	#	Variable	B	SE B	ß	t	Sig
1	1	Perceptual salience	.01	.00	.77	2.75	< .001
	4	Frequency	.24	.09	.69	2.63	< .01
	2	Morphophonological regularity	-.12	.07	-.34	-1.53	.12
	3	Syntactic Category	-.39	.27	-.52	-1.41	.15
2	5	L2 Proficiency	.83	.11	.46	7.38	< .001
	6	L1	.08	.07	.12	1.14	.25
	7	L1-L2 Mismatch	.03	.09	.03	.30	.76

Note. SE = standard error. Method: forced-entry in SPSS, n = 160. For model 1, R = .38. R^2 = .15. For model 2, R = .58. R^2 = .34. Semantic complexity was excluded by SPSS 19 for failing collinearity requirement.

Variables with significance value lower than .05 in the last column, perceptual salience, frequency, and L2 proficiency, are the ones meaningful to the regression and make the best contribution to the

respective model as expressed by their β value. Comparing results with those of L2 English shows perceptual salience, frequency, and L2 proficiency are the only variables to explain substantial amounts of variance cross-linguistically when all other variables are held constant.

Finally, model 1 for L2 Italian yields worse explained variance compared to L2 English in both models (R^2 nearly 15 vs. 17% for model 1, R^2 = 34 vs. 40% for model 2). The improvement when adding the new non-linguistic variables is slightly worse in L2 Italian than L2 English insofar as explained variance jumps by 20% in the former case, in contrast to 23% in the latter. These results, therefore, only partially confirm the second prediction that salience factors have predictive validity in L2s other than English, as only 2 of the 5 original predictors, frequency and perceptual salience, were found to be significant.

7. Discussion

The first aim of this study was to ascertain that five salience factors, frequency, perceptual salience, morphophonological regularity, semantic complexity, and syntactic category were able to explain large amounts of variance in accuracy scores elicited via written tests, rather than oral production. Indeed it was found 2 out of 5 salience-related variables, perceptual salience and frequency, were meaningful to the L2 English and L2 Italian regressions, consistent with G&D (2001) and Ellis (2006). Triangulating with very different tests compared to the original morpheme order studies meta-analysed by G&D allowed us to exclude, with a certain degree of confidence, that the predictive validity of the two factors is due to a task effect (i.e. only inferable from oral production).

What does the finding that perceptual salience and frequency determine accuracy across a variety of L2 English studies and task types imply for L2 acquisition? In general, when knowledge of L2 morphology is found to differ between native and non-native speakers under certain task conditions (e.g. oral production, written production, grammaticality judgment, or comprehension only), it is associated with differences in processing or computational mechanisms. This contrasts sharply with results of experiments in which native and non-native knowledge of L2 morphology is found to differ systematically across different types of test, as these imply more profound differences at the level of representation of knowledge. (Hopp 2012, 2013; McCarthy 2008; Snape et al. 2009). As Hopp (2012: 2) puts it:

> In general terms, (second) language learners take four steps in acquiring L2 inflectional morphology. First, they need to acquire morphological forms and store them in the mental lexicon; second, they need to map the forms to the correct interpretations and identify the grammatical contexts of their occurrence; third, they need to retrieve the correct forms in the requisite grammatical contexts and, last, they need to be able to spell out the forms correctly in production. The first two steps comprise the acquisition of grammatical knowledge of the morphophonology and the morphosyntax of inflection, respectively; the latter two steps implicate lexical retrieval and the real-time processing of inflection.

If perceptual salience and frequency only affected production of *-s* and *-ed* (and hence knowledge of the associated morphosyntactic features 3SG, present tense, and past participle) in the oral modality, then such factors affect the latter two steps which implicate lexical retrieval and processing. This study, instead, shows that this is not the case. Salience-related factors additionally affect accuracy in written production, complementing an already large pool of studies showing native and non-native knowledge of L2 morphology differ in oral production. Therefore, the effects of perceptual salience and frequency are not limited to the latter two stages described by Hopp (2012) but are relevant to the first two stages with consequences for the abstract knowledge or representation of morphological knowledge in L2 learners as well. Future studies of the L2 acquisition of L2 morphology, therefore, should continue to operationalise morphological forms in terms of perceptual salience and frequency.

To expand further on the predictive validity of the five salience factors, morphophonological regularity and semantic complexity were also found to be a significant predictor, albeit for the L2 English regression only. The former was operationalised as the sum of the numerical value of two sub-factors, allomorphy and homophony with other grammatical morphemes. In consequence, this leads to the prediction allomorphy and homophony of verbal inflection should be highly predictive of accuracy in L2s typologically similar to English, such as Dutch and German, whose morphology is also stem-based. The prediction is the higher the number of allomorphs and homophony, the lower the accuracy score. The latter variable was predicted to correlate negatively to the average accuracy score. Instead, the opposite relationship was found: the higher the semantic complexity score, the higher the accuracy score. In contrast with G&D's study where the same variable was found to be, in line with prediction, inversely

proportional to accuracy scores, this finding suggests that verb morphemes are better perceived if they encode several morphosyntactic features such as tense, person, number, and gender. When semantic complexity is extemporaneously extended to a heterogeneous set of morphemes as in G&D, though, the relationship reverses.

The next prediction, that salience would play a significant role in the acquisition of a closed set, was only partially confirmed. Semantic complexity, syntactic category, and morphophonological regularity were not as meaningful to explaining variance in accuracy of verb morphemes as they were for the set in G&D which comprised nominal, verbal, and determiner-related morphemes. It is, therefore, plausible that the interplay of salience factors is strictly dependent upon the class of morphemes investigated. More interestingly, certain inflectional morphemes occupied the same position in the L2 English (Table 5) and L2 Italian (Table 8) accuracy ranks, indicating use is contingent upon the type of inflection. In particular, this study found that across two typologically different L2s whose inflectional morphemes differ especially in respect of their salience properties in the input, learners were more accurate in use of a semantically empty verb, the copula, followed by the most morphologically productive 3SG present tense, perceived as /z/ in English but /a/ in Italian, and past participle forms, perceived as /t/ in English but /ato/ in Italian, ending with the less morphologically productive 3SG present tense forms, perceived as /s/ in English but /e/ in the Italian. Thus, free forms are used more accurately than bound forms as previously found by Zobl & Liceras (1994).

Another important contribution of this study relates to the differences in mean accuracy of 3SG when it is perceived as /s/ and /z/. Research on the acquisition of L2 English morphology hitherto has not focused on differences in production of 3SG morphology by perception of allomorphs principally because aspects of salience of the input have not received equal attention to other factors. This study has shown that lexical storage of the abstract morphological features 3SG and their phonological realisation (/s/ and /z/) are loosely connected to accuracy rates in written production: L2 learners have far less difficulty affixing 3SG *-s* when the associated lexical item accessed is perceived as /z/ rather than /s/. Furthermore, as shown in Table 7, the token frequency and transitivity of the predicates tested did not appear to be directly associated with the their ranked accuracy. However, this study did not set out to investigate linguistic factors beyond the five factors related to salience, L1-L2 mismatch, 'language background' (L1), and L2 grammatical proficiency; based

on its design, the study cannot conclusively rule out the importance of other linguistic variables such as token frequency and transitivity but offers suggestive evidence that differences in the accuracy levels for 3SG *-s* are more likely to be associated with its variant allomorphic realisations.

One variable not considered in the study which is open to future investigation is word familiarity of the predicates. The null hypothesis in this study has been that familiarity is directly proportional to token frequency (Tanaka-Ishii & Terada 2011) implying, in the context of the present study, that it is unrelated to accuracy levels of the 3SG *-s* morpheme. If, however, familiarity were assumed otherwise, future psycholinguistic studies may be designed to classify predicates by the participant's familiarity with them and set out to look for a relationship with accuracy rates of L2 morphology.

Another variable not considered are L1-L2 phonotactic differences in the syllables composing the verb stems tested. To exemplify, it may be that the accommodation of phonotactic structures such as English *go/sell* and *check/reflect* in French and Spanish, as well as Italian *andare/controllare* and *vendere/riflettere* in English and Spanish plays a facilitative role in use of affixes. This hypothesis, however, would predict no substantial differences between average accuracy for the four stems in the French and Spanish groups since all four stems are possible phonotactically in either L1 (e.g. *check* as *checar* in Spanish and *chèque* in French) which was not confirmed by the data as shown in Table 6. In the Italian data, the same hypothesis would predict both the Spanish and English to score lower for *riflettere* than other stems since the combination of sounds or orthographic forms in the root *riflet-* is part of neither the Spanish nor the English lexical inventory. The prediction, in fact, is borne out by the data given the English and Spanish average scores for the 3SG of the verb *riflettere* is the lowest across the board (gap-fill: English *X-bar* = .78, Spanish *X-bar* = .58; sentence completion: English *X-bar* = .40, Spanish *X-bar* = .35). These results, thus, recommend future research take into account L1-L2 phonotactic differences in the structure of verb roots.

The L2 Italian segment of the study has shown that L2 learners achieve far higher levels of accuracy when the lexical item for 3SG to be retrieved is attached to roots of the most productive morphological class ending in -are. As shown in Table 4, stems formed by a root plus a 3SG *-a* or 3PL *-ano* suffix of the first class also have very high token frequency. Thus, it is possible that L2 learners are more accurate in use of verbs of the first rather than the second class because the former are more frequent in the input.

As for the question of cross-linguistic plausibility, the symmetry in good predictors found across L2s suggests perceptual salience and frequency are the two most important top down salience factors in the acquisition of L2 verb morphology. The symmetry suggests the impact of differences in the intrinsic morphological processes between the morphology of the two languages, namely the type of conditioned allomorphy and direction of affixation, is minimal. More research on L2s such as Dutch and German, whose morphemes are subject to different computational processes than Italian, is necessary to corroborate our finding that accuracy of use is largely unaffected by the morphological processes involved. Such line of investigation is also beneficial to further validate the predictive validity of salience as a universal factor.

Lastly, when three new variables were explored by adding them to the regression: (1) L2 proficiency turned out to be a significant predictor of variance in the DV for both L2 English and L2 Italian; (2) increases were found in the model fit, particularly for the L2 English regression which yielded a 23% improvement from the model including the salience factors only; (3) language background was nearly significant in the L2 Italian regression. The present study enabled us to look at the individual contribution of one crucial factor disregarded in the literature on morpheme order studies: L2 proficiency. This factor turned out to be the most powerful predictor of accuracy among the learner's L1, L1-L2 mismatch, and any of the salience factors, across L2s. However, in this study it was shown that the factor learner's L1 was significant in the L2 English regression. A reviewer aptly points out this finding may be reconcilable with the relative typological difference between the L1s and the L2s in terms of verbal morphology, consistent with previous reviews of the morpheme order studies (Kwon 2005; Luk & Shirai 2009). In other words, language background nearly turned out to play a significant role in predicting accuracy scores of L2 English inflection because the L1s, French and Spanish, are both Romance languages typologically related to each other but not the L2. Such configuration contrasted with L2 Italian where language background did not reach statistical significance and the L1s were less similar to each other insofar as one belongs to the Romance family, the other to the Germanic. In practical terms, the results of the English segment of the study suggest accuracy scores are higher if the L1 value is 1: Table 7 reported a *B* value of .27, when the independent variable was L1 Spanish.

8. Conclusion

This study was an extension test for the five top down salience factors of the input previously found by G&D (2001) to be strongly predictive of accuracy scores in use of a heterogeneous set of L2 grammatical morphemes by L2 English learners. It investigated the extent to which: (1) the predictive power of salience was not an artifact of the testing via oral production in the morpheme order studies; (2) salience could predict accuracy in use of a narrow set of verb inflections and is cross-linguistically valid; (3) three new variables could predict residual variance. In this way, the study also sought to address limitations of G&D's design, as well as other criticisms of the morpheme order studies. Multiple regression analyses showed two factors, perceptual salience and frequency, were good predictors of accuracy scores across L2s which differ especially in relation to salience, English and Italian. Moreover, these two factors, previously believed to impact on accuracy of use of morphology in oral production, were found to also impact on written production tests where L2 learners have more time to access the appropriate lexical forms. In line with previous research in the field, this finding confirms perceptual salience and frequency are of profound importance to the acquisition of the abstract morphosyntactic features associated to verbal morphology. The interplay of salience factors, however, is dependent upon the class of morphemes investigated because they lose their predictive validity if only verb inflection is considered.

L2 learners of inflectionally rich, as well as impoverished L2s, are more likely to use target-like inflection if producing a semantically empty verbs, namely the copula, followed by the most morphologically productive 3SG present tense inflection, the past participle form, ending with the less morphologically productive 3SG present tense forms. It is not, however, clear whether copula forms were used more accurately due to their semantic content or morphological unboundedness property. In this respect, decreased semantic complexity and free morphological form appear to conspire to explain accuracy. Testing three new variables, L2 grammatical proficiency, language background, and L1-L2 mismatch for the type of verb affixation, showed L2 grammatical proficiency to be a significant predictor. Although language background was not a significant predictor cross-linguistically, a significant effect was found in the L2 English regression where the L1 Spanish participants had an advantage over the French, underscoring the relevance of this variable to future studies.

Notes

[1] Evidence in favor of a disjunction between sonority of voiced in contrast to voiceless obstruents is inconclusive and has lead to different hierarchies in the literature (Parker, 2002). Depending on the approach to voicing endorsed, permutations will vary. Another aspect affecting the relative order of phonemes in the hierarchy is tied to a view to arrive at a scale with universal validity versus one that is language-specific. Nevertheless, we have decided to adopt Parker's hierarchy because it suits the languages experimented with in our study.

[2] The choice of child-directed speech is somewhat problematic in G&D (2001) as it deviates from adult-directed speech of the kind adult L2 learners are exposed to. Alongside that, adult child-directed speech is often simplified compared to adult-directed speech. To keep with the original study, however, I apply the same methodology herein.

[3] Although scores for *-ato* may be computed for the *-at* suffix, we wanted the scoring to be consistent across all the other morphemes which are indivisible (*-a*, *-e*, *è*, *sono*). Moreover, this approach would consider TMA in isolation from gender and number, which is undesirable since morphemes such as *-s* were not examined for person and number in isolation of features such as tense.

[4] Because English 3SG *-s* and its allomorphs are homophonous with other morphemes such as plural *-s*, the genitive *-s*, and their relative allomorphs, even in oral production one can never be fully sure whether the lexical form being scored is being matched to the appropriate morpho-syntactic features. To illustrate, if a participant produces the verb *speak-s*, pronounced as /spi:ks/, in a 3SG context, there is no reliable way to exclude the possibility that the /s/ sound is mapping 3SG verb features as opposed to the plural or possessive marker of a noun. In the presence of homophony, there is no a priori way to distinguish when the inflection used is verbal or nominal. This, however, is not the case for Italian where the 3SG allomorphs tested are not homophonous to any nominal forms.

[5] The tests were modelled after recommendations in Ellis (2006: 434) for distinguishing items that tap into implicit versus explicit knowledge of grammatical properties while the timing of the tests was set according to the method proposed in Ellis (2005: 156).

[6] A reviewer suggests providing the L2 English learners with a prompt in the bare infinitive form, in contrast to the L2 Italian who are given an inflected infinitive may result in bias. However, there are a number of reasons why the diversity in prompt does not represent a threat to the test design. English speaking children and L2 English learners apply different morphological processes and develop verb morphology in very different ways to Italian speaking children and L2 Italian learners. English speaking children and child/adult L2 speakers overuse bare (infinitive) forms in contexts where other forms inflected for tense and agreement are required in early stages of language development (e.g. *reflect* for *reflects* or *reflected*) (see Rizzi, 1993 and Wexler, 1994 for L1; Prévost & White, 2000 for L2). In other words, there is robust evidence that native and non-native English grammars are characterised by morphological omission. On the other hand, Italian speaking children and child/adult L2 speakers, though much more proficient than their English counterparts in knowledge and use of verb inflection at early stages of language development, overuse a stem form most frequently homophonous to either the 3SG present indicative form (Guasti 2002 for L1; Giacalone-Ramat 2004 and Rodgers 2011 for L2) or the imperative (Salustri & Hyams 2002 for L1; Romano 2014 for L2). This default form, thus, reflects morphological errors in native and non-native Italian are fairly different from English because they result in substitution rather than omission.

[7] The timing for all tests was established during the pilot stage of the experiment by doubling the time needed by native speakers to complete the tests. See also Ellis (2005) for this technique.

[8] In the dataset, there were only two such cases in L2 English and one in L2 Italian. By creating obligatory contexts for 3SG and 3PL we tease apart incorrect anaphora resolution between the target subject and a preceding topic and SV agreement. In other words, participants to the L2 English study are still capable of supplying correct inflection for the verb even if the agreeing subject is incorrectly omitted (e.g. by Spanish speakers whose grammar allows null subjects) or not agreeing in features with the referent antecedent. An alternative view is that subject triggers agreement on the verb which is an agreement target. Under this approach, cases like *reflects* are uninformative (and should be excluded from analysis), since one cannot tell what is triggering agreement on the verb; cases like *they reflects* are not really cases of an incorrectly inflected verb; and cases like *he reflect* are cases of omission. Coding the data in this manner does not change the results. I thank a reviewer for this suggestion.

Bibliographical References

Ashworth, Julie & Clark, John 2005. *Stepping Stones 2*. New York: Thomas Nelson & Sons.

Awh, Edward; Belopolsky, Artem V. & Theeuwes, Jan 2005. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Science* 16 (8). 437-443.

Bailey, Nathalie; Madden, Carolyn & Krashen, Stephen 1974. Is there a "natural sequence" in adult second language learning? *Language Learning* 24. 235-243.

Bresnan, Joan 2001. *Lexical-functional syntax*. Malden, Mass.: Blackwell.

Brown, Roger 1973. *A First language: The Early Stages*. Cambridge, Mass.: Harvard University Press.

Chiarcos, Christian; Claus, Berry & Grabski, Michael 2011. S*alience: Multidisciplinary perspectives on its function in discourse*. New York, Berlin: de Gruyter.

Dalrymple, Mary 2001. *Lexical Functional Grammar*. New York: Academic Press.

De Mauro, Tullio; Mancini, Federico, Vedovelli, Massimo & Voghera, Miriam 1993. *Lessico di frequenza dell'italiano parlato*. Milan: Etas Libri.

Docherty, Gerard J. & Foulkes, Paul 2014. An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua* 142. 42-56.

Dulay, Heidi & Burt, Marina 1974. Natural Sequences in Child Second Language Acquisition. *Language Learning* 24. 37-53.

Ellis, Nick & Larsen-Freeman, Diane E. 2006. Language emergence: Implications for applied linguistics. *Applied Linguistics* 27 (4). 558-589.

Ellis, Nick 2002. Frequency Effects in Language Processing. *Studies in Second Language Acquisition* 24. 143-188.

Ellis, Nick 2006. Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing,

blocking, and perceptual learning. *Applied Linguistics* 27 (2). 1-31.
Ellis, Rod 2005. Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27. 141-172.
Ellis, Rod 2006. Modelling Learning Difficulty and Second Language Proficiency: The Differential Contributions of Implicit and Explicit Knowledge. *Applied Linguistics* 27 (3). 431-463.
Gass, Susan & Mackey, Alison 2002. Frequency effects and second language acquisition: A complex picture? *Studies in Second Language Acquisition* 24 (2). 249-260.
Giacalone-Ramat, Anna 2004. *Verso l'Italiano: Percorsi e strategie di acquisizione*. Roma: Carocci.
Giraudo, Helene & Dal Maso, Serena. 2015. Suffix perceptual salience in morphological processing: evidence from Italian. *Proceedings of the NetWordS Final Conference, March 30-April 1.* Pisa, Italy.
Goldschneider, Jennifer M. & DeKeyser, Robert. M. 2001. Explaining the "Natural Order of L2 Morpheme Acquisition" in English: a Meta-analysis of Multiple Determinants. *Language Learning* 55. 27-62.
Hawkins, Roger 2001. *Second Language Syntax*. Oxford: Blackwell Publishing.
Hopp, Holger 2012. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29 (1). 33-56.
Hopp, Holger 2013. The development of L2 morphology. *Second Language Research* 29 (3). 1-5.
Ibánez, Ursula 2013. *A morpheme order study based on an EFL learner corpus: A focus on the Dual Mechanism*. Ph.D dissertation. University of Granada, Granada, Spain.
Jany, Carmen; Gordon, Matthew, Nash, Carlos M. & Takara, Nobutaka 2007. *How Universal is the Sonority Hierarchy?: a cross-linguistic acoustic study. Proceedings of the 16th International Congress of Phonetic Sciences* 2007, Saarbrücken, 6-10 August 2007.
Järvikivi, Juhani; Bertram, Raymond & Niemi, Jussi 2006. Affixal salience and the processing of derivational morphology: The role of suffix allomorphy. *Language and Cognitive Processes* 21 (4). 394-431.
Krashen, Stephen; Sferlazza, Victoria, Feldman, Lorna & Fathman, Ann K. 1976. Adult performance on the SLOPE test: More evidence for a natural sequence in adult second language acquisition. *Language Learning* 26. 145-151.
Kwon, Eun-Young 2005. The "Natural Order" of Morpheme Acquisition: A Historical Survey and Discussion of Three Putative Determinants. *Columbia University Working Papers in TESOL & Applied Linguistics* 5. 1-21.
Langman, Juliette & Bayley, Robert 2002. The acquisition of verbal morphology by Chinese learners of Hungarian. *Language Variation and Change* 14. 55–77.
Lardiere, Donna 2006. Attainment and acquirability in second language acquisition. *Second Language Research* 22 (3). 239-242.

Lardiere, Donna 2007. *Ultimate Attainment in Second Language Acquisition: A Case Study*. Mahwah, NJ: Lawrence Erlbaum Associates.
Larsen-Freeman, Diane 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9. 409-419.
Larsen-Freeman, Diane 1976. An explanation for the morpheme acquisition order of second language learners. *Language Learning* 26. 125-134.
Luk, Zoe P. & Shirai, Yasuhiro 2009. Is the Acquisition Order of Grammatical Morphemes Impervious to L1 Knowledge? Evidence From the Acquisition of Plural -s, Articles, and Possessive 's. *Language Learning* 59 (4). 721-754.
MacWhinney, Brian 2000. *The CHILDES Project: Tools for analysing talk*. Mahwah, NJ: Lawrence Erlbaum Associates.
McCarthy, Corinne 2008. Morphological variability in the comprehension of agreement: an argument for representation over computation. *Second Language Research* 24 (4). 459-486.
Pica, Teresa 1983. Adult Acquisition of English as a Second Language under different conditions of exposure. *Language Learning* 33. 465-497.
Pienemann, Manfred 1998. *Language Processing and Second Language Development: Processability Theory*. Amsterdam: Benjamins.
Poeppel, Douglas & Wexler, Kenneth 1993. The full competence hypothesis of clause structure in early German. *Language* 69. 1-33.
Prévost, Philippe & White, Lydia 2000. Missing surface inflection or impairment in second language acquisition. Evidence from tense and agreement. *Second Language Research* 16. 103-133.
Rodgers, Debora 2011. The Acquisition of Verbal Agreement in Instructed Italian L2A. *L2 Journal* 3. 93-115.
Romano, Francesco 2014. *The role of Universal Grammar, the L1, and the input in the development of L2 inflection*. Ph.D dissertation. University of Essex, Colchester, UK.
Rosado, Luis A. 1986. *The role of the Monitor in the acquisition sequence of twelve English morphemes by adult English as a second language learners from four different language groups.* Ph.D. dissertation. Texas A & I University, Kingsville.
Salustri, Manola, & Hyams, Nina 2002. *Is There an Analogue to the RI Stage in the Null Subject Languages? Paper presented at the 27th Annual Boston University Conference on Language Development*. Boston, Massachusetts.
Snape, Neal; Yan-kit, Ingrid Leung & Sharwood-Smith, Michael 2009. *Representational Deficits in SLA: Studies in honor of Roger Hawkins*. Amsterdam: John Benjamins Publishing.
Zobl, Helmut & Liceras, Juana 1994. Functional categories and acquisition orders. *Language Learning* 44. 159–80.