University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Language Acquisition Work by Anne Vainikka   Anne Vainikka Works Site

2018

Solving the mystery of the missing surface inflections
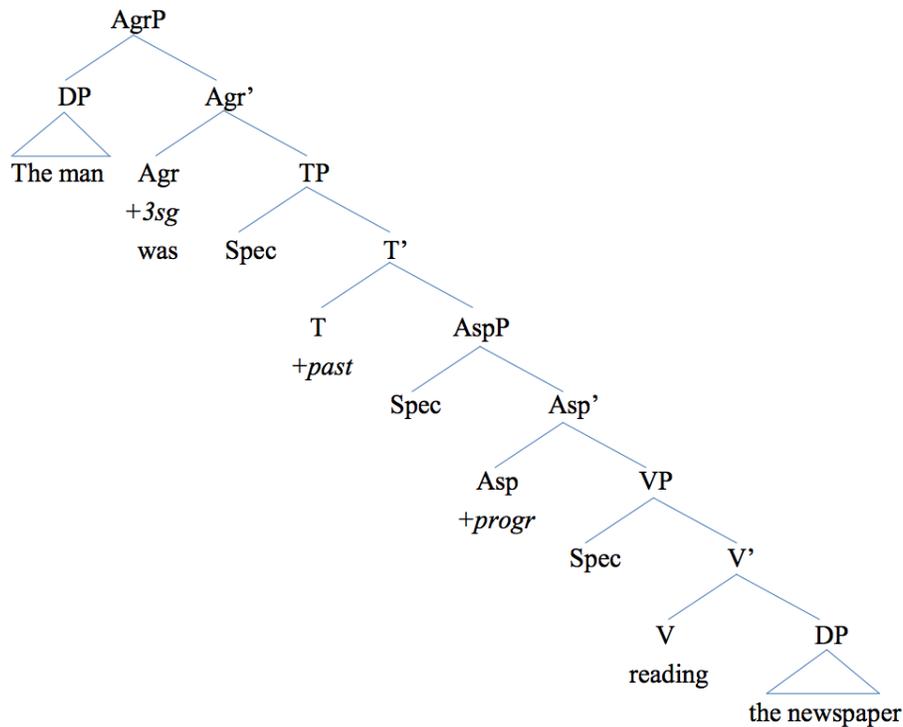
Anne Marjatta Vainikka

Introduction

Second language acquisition researchers and teachers have long suspected that even though learners with plenty of target language exposure vary in their production of inflectional morphemes such as regular past tense and third person singular –s (the only agreement suffix in English), they do know these forms. Researchers ask why learners don't invariably produce these forms if they know them. But do they really know them? To answer these questions, we need to ask what it means to "know" (have acquired) an inflectional morpheme. Under current thinking, a learner's knowledge of morphological forms marking tense and agreement is connected to the abstract syntactic features they represent. As noted in other chapters in this volume (e.g. chapters 2 and 4 by Caink, and by Gil *et al*. respectively), in generative, Chomskyan, linguistics, human languages share a common core of syntactic principles, and differences among languages revolve around language-specific mental lexicons. The lexicon of a language contains the semantic, phonological, morphological and syntactic properties of lexical categories such as verbs, nouns, adjectives and prepositions and **functional categories s**uch as negation, tense, aspect and agreement.

Under more recent ideas about syntax known as Minimalism (Chomsky, 1993, 1995, 1998, 2001), the lexicon contains functional morphemes with their **formal features s**uch as [+/- past] which determine a language's specific syntactic structure or its projections. Lexical projections include VP whose head is a V (verb) and whose arguments (e.g. a direct object) either precede or follow the head (Farsi and Japanese vs. Arabic, Chinese or English). Under **X'-Theory (**Jackendoff, 1977), every head (X) in a sentence projects a phrase (XP) in a hierarchical or tree-structure as illustrated in (1) below, in an English example. In the VP, the lowest projection, relations between the verb's arguments – e.g. subject and object – are established. The higher

verbal projections in the tree are functional: AspP (aspect phrase), TP (tense phrase) and AgrP (agreement phrase).[i] DP (determiner phrase) is a nominal projection which in English includes functional morphemes such as articles. For further details on syntactic structure, see Caink's chapter 2 at the start of this volume.

(1)

```
                        AgrP
                       /    \
                     DP      Agr'
                    /       /    \
              The man     Agr     TP
                          +3sg   /  \
                          was  Spec   T'
                                     /  \
                                    T    AspP
                                  +past  /   \
                                       Spec   Asp'
                                             /    \
                                           Asp      VP
                                          +progr   /  \
                                                 Spec   V'
                                                       /   \
                                                      V      DP
                                                  reading
                                                         the newspaper
```

Morphosyntax refers to the position of constituents in a sentence (i.e. word order) in relation to functional morphology in a sentence. Acquisition is a process of receiving input about the nature and type of features in that language. When second language (L2) learners inconsistently produce –ed or –s, researchers ask whether their **interlanguage** contains AgrP or a TP, taking one of two positions: (1) the learner's interlanguage contains these projections, but the learner does not always produce the morphemes due to the pressure of producing something during a conversation or for phonological reasons, i.e. the learner's first language (L1) disallows the consonant clusters which

2

result from suffixation of –ed or –s to a word-final consonant as in 'grabbed' or 'drinks'; (2) the learner's interlanguage has no AgrP or TP. This chapter argues for view (2) on the basis of evidence from a set of different experiments, which we briefly discuss, to explore whether learners have acquired the relevant syntax. We start with what we know about L1 acquisition given the commonalities we assume between L1 and L2 acquisition.

What's missing in the morphosyntax of children?

Children's first multi-word utterances reveal their early grasp of VP word order. At age one year and ten months – 1;10 – little German learner Meike produced the complement before non-finite *gehen* 'go' as in (2), i.e. a head-final German VP. But this is not adult German: the verb follows the subject and precedes the preposition phrase: *Ich gehe nach Hause* 'I go home' or *Ich will nach Hause gehen* 'I want to go home' (Mills, 1985). Adam at age 2;3 placed the complement 'doggie' after 'give', as in (3), in keeping with the head-initial English VP (Brown, 1973).

(2)   Hause gehen

       home go-INF

(3)   Give doggie.

Children's utterances lack certain elements: in both examples, the subject is missing, and in the German one, the verb is non-finite (in English lack of agreement is more difficult to ascertain).[ii] Where an auxiliary verb is required, it is clearer that L1 English children omit functional elements, e.g. in typical Wh-questions at early stage (Klima & Bellugi, 1966):

(4)   a. Where go?

       b. What doing?

Children learning other European languages show similar omissions .[iii] Similar examples from children at the early stages show omission of functional elements required in adult speech including bound morphemes and free morphemes such as auxiliary and modal verbs.[iv] Early views were that this initial system involved roles such as Agent and Patient and that children moved in a discontinuous way from a semantic system to a syntactic one. Alternatively, children's earliest system could be syntactic, constrained by Universal Grammar (UG, Chomsky 1981) whose syntactic principles require finite clauses to have subjects and inflected verbs (the specifier of AgrP assigns nominative case to a subject. Innately-driven language acquisition has long been supported by studies revealing strikingly uniform developmental patterns across children. These include Brown's (1973) longitudinal study of three children learning English and de Villiers and deVilliers' (1973) cross-sectional replication with 21 children. It's been noted that the likelihood of these common patterns being due to chance is practically zero (Laurence & Margolis, 2001; see also Stromswold, 2000).

Under UG-driven acquisition, adult languages share functional categories, and principles relate to a common syntactic structure for all languages (apart from variation expressed as parameters, e.g. for word order). Even though children's tense and agreement suffixes may be missing from their utterances, their syntax nonetheless contains TP and AgrP (this view is known as Strong Continuity; see Boser, Lust, Santelmann and Whitman, 1992; Hyams, 1992, 2007; Lust, 2006; Poeppel and Wexler, 1993; Wexler, 2004). But if children's syntax contains TP and AgrP, why do they omit similar suffixes as well as subjects? Researchers have proposed that the child truncates their underlying syntax in a way not possible for adults. Children do so until the principle of 'categorical uniformity' matures around age three.[v]

Maturation of categorical uniformity would be an ideal solution if older, already mature second language learners did not produce similar utterances. But they do, as we will see below. An alternative Weak Continuity view holds that UG provides the tools for the child to select the

specific subset for their language from a universal set of functional categories. Vainikka and Young-Scholten (2007) discuss an analysis by Culicover and Jackendoff (2005) of the historical development of ideas in generative syntax and how these ideas should but only sometimes do relate to children's language acquisition.

In his seminal contributions to children's acquisition of English, Radford (1988, 1990, 1995) proposed that children's early system is syntactic, but it lacks functional projections. The examples in (2), (3) and (4) and footnote 4 all involve a **minimal tree,** just a VP without AgrP or TP. Radford also took a maturation approach whereby all functional categories mature together (see Clahsen, 1991; Clahsen and Penke, 1992; Clahsen, Eisenbeiss & Vainikka, 1994 for similar ideas for German). A more refined, non-maturationist approach involves the child building syntactic structure, one functional projection at a time. Vainikka (1993/4) considered data from children from CHILDES.[vi] One child, Nina, at age 1;11 to 2;0 produced utterances with incorrect *my*-subjects along with non-finite verbs, as in (5a), and produced nearly no modals, no auxiliary verbs, past or agreement suffixes, embedded clauses or inverted questions. The absence of all of these indicates a VP-only grammar. At 2;1, Nina produced numerous nominative *I*-subjects along with previously omitted verbal morphology as in (5b). A window on Nina's immature morphosyntax is provided by *my*-subjects in Wh-constructions as in (5c) which suggest no functional projections above AgrP for production of adult-like questions requiring the higher projection CP (complementizer phrase) for question formation, relative clauses and embedded clauses. The subject in Wh-constructions remains in VP, but cannot get nominative case for 'I'.

(5)   a. My make a house.   [Nina 2;0]

        b. I don't break 'em   [Nina 2;1]

        c. Look what my got.   [Nina 2;3]

A later longitudinal study by Ingham (1998) suggests that between 2;6 and 2;9 children

acquiring English acquire TP following the tree in (1), building syntactic structure from the bottom up, from the VP. There was not yet evidence for an AgrP: the child's utterances did not show contrasts for case (on pronouns) or for agreement.

What's missing in the morphosyntax of L2 learners?

Researchers also disagree about whether second language learners at early stages have a full syntactic tree or a minimal VP tree. For those who assume UG-driven acquisition across the lifespan, the second language learner also brings their knowledge of their first language to the task of L2 acquisition. There is a long tradition – dating at least back to Lado's (1957) Contrastive Analysis Hypothesis (CAH) – of considering how the L1 shapes L2 acquisition. In the 1970s, the CAH was rejected on various conceptual and empirical grounds, including evidence from studies of L2 English revealing a common acquisition order. That is, not only child but also adult L2 learners regardless of their L1, displayed in cross-sectional studies common accuracy orders in production of inflectional morphology, and common stages in acquisition of questions and negation. L2 learners' early production showed omission of functional morphemes and their systematic emergence (Bailey, Madden & Krashen, 1974; Dulay & Burt, 1974). Despite rejection of L1 influence, researchers continued to investigate how the L1 influences L2 acquisition, because (1) omission of functional elements persists and (2) its rate varies across L1 groups. This has meant on-going search for the source of missing inflections; for L2 English see White (2003), and Haznedar (2001) for L1 Turkish; see Ionin and Wexler (2002) for L1 Russian and Lardière (1998, 2007) for L1 Chinese; see Prévost and White (2000 a, b) for various L1 learners of L2 French or L2 German; see Vainikka and Young-Scholten (2011) on various L1 learners of L2 German. To account for missing inflections, there are three general approaches: Strong Continuity, Weak Continuity and processing.

Full Competence

Full Competence is the Strong Continuity approach in L2 acquisition, expressed by Schwartz and Sprouse's (1994; 1996) Full Transfer/Full Access Hypothesis. It assumes the L2 learner's initial state of knowledge includes their L1 syntax, in addition to UG. The solution to the mystery of missing inflections is the Missing Surface Inflection Hypothesis: learners' interlanguage grammar is a full syntactic tree, but inflectional morphology has to be learned and the phonology involved also has to be mastered. Others argue that learners cannot acquire the relevant syntax if tense and agreement are not part of their L1 syntax; their interlanguage syntax is 'defective'.

The Missing Surface Inflection Hypothesis (MSIH)

One piece of evidence supporting the MSIH (Epstein, Flynn & Martohardjono, 1998; Haznedar & Schwartz, 1997; Lardière, 1998) is the second language learners' production of pronominal subjects in correct, nominative case, regardless of whether learners produce non-finite or finite verb forms in the same utterance. This indicates that learners have an AgrP (and a TP) since AgrP is syntactically required for prominal subjects to get nominate case. Why then, should learners sometimes omit the inflection involved in AgrP if they have acquired the syntax? To explain why learners variably produce the relevant suffix, those who support MSIH apply **Distributed Morphology** (Halle & Marantz, 1993). The explanation is as follows: in attempting to produce morphology, the speaker has several possibilities. Let's consider the agreement paradigm in a richly inflected language such as Spanish with its separate suffixes on the verb which mark first, second and third person in both singular and plural. When the language learner has to match a suffix to the right person and the right number, there is competition, and what we refer to as the least specified suffix will win. This could be a form without a suffix at all. Distributed Morphology accounts for what researchers often observe during acquisition: the learner makes systematic errors of omission of inflection, e.g. the learner uses non-finite or stem forms as in (2) and (3) above rather

than using the wrong inflection (e.g. a second person suffix for third person) (so called errors of commission). This variable production of inflection by the learner is the result of problems with sufficiently rapid processing (there isn't enough time for the learner to retrieve the correct suffix). This seems to be a problem for older L2 learners (i.e., adults). O'Grady (2006) and his chapter 3 in this volume describes a similar approach, although he takes a general cognition rather than a UG approach.

Variable production of inflection could be due to the influence of the learner's L1 phonology. Goad, White and Steele (2003), in their Prosodic Transfer Hypothesis, considered data showing that the L2 English production of speakers of Chinese, which allows no consonant clusters at the beginnings or ends of words, depended on mode of production. The speaker in Lardière's (2003, 2007) study produced past tense 78% of the time in written production, but only 35% of the time in oral production. That this is a phonological problem is shown in her more successful oral production of irregular forms without clusters such as 'saw', and in omission of consonants in words with single morphemes with final consonant clusters such as 'strict'. According to O'Grady (2006), inflection marking is carried out by the learner's general cognitive processor and it works linearly, in the order that clauses are produced or processed.

Defective syntax

Other accounts for variable production of inflection are the Failed Functional Features Hypothesis (Hawkins & Chan, 1997) and its successors, the Local Impairment Hypothesis (Beck, 1998), the Representational Deficit Hypothesis (Hawkins, 2003) and the Interpretability Hypothesis (Tsimpli & Dimitrakopoulou, 2007). We refer to these as defective syntax hypotheses and all of them hold that older L2 learners can only acquire those features present in their L1; access to Universal Grammar is partial. Hawkins and Liszka (2003) found lower production of English past -ed by speakers of [-tense] Chinese (63%) than by speakers of [+tense] Japanese (92%). But this raises the question of how Chinese speakers come to mark tense at all? It is possible

that we observe the marking of tense because the evidence of variable production of past tense comes from instructed learners, and their tense marking may a result of what they have learned in the classroom rather than what they have acquired (Krashen, 1985; Schwartz, 1993). It is also feasible that Chinese learners are using a verbal features in their L1, namely their aspect feature, as tense in English, a view expressed in Lardière's (2008) Feature Reassembly Hypothesis. Under this hypothesis, learners recruit their L1 features and, if required, assemble them differently for the language they are in the process of acquiring.

Weak Continuity in L2 acquisition

Another approach to missing morphology is a Weak Continuity approach under which the learner starts with a syntactic structure which lacks functional projections, i.e. a VP projection, the base of the tree shown in (1) above. As input is received, the learner then develops functional projections in succession. Vainikka and Young-Scholten's (2011) Organic Grammar holds that learners start with a minimal tree and then build functional structure. The theory dates back to Vainikka and Young-Scholten's (1994) **Minimal Trees Hypothesis** in which they argued on the basis of data from uninstructed adult immigrants in Germany that the learner's 'initial state' when they start acquiring an L2 is a bare VP. This VP is initially transferred from their L1; that is, it has the word order of the VP in their native language. Under the MSIH, the assumption is that utterances with non-finite verb forms involve full syntax, so subjects are marked with nominative case by AgrP. We find the same L2 early multiword utterances which, however, lack subjects as we find in L1 acquisition. (L2 beginners also produce the same sort of single-word and verb-less utterances children learning their first language produce; see Myles, 2005). An example in the L2 acquisition of English comes from Yamada-Yamamoto (1993) of her Japanese-speaking son's earliest stages of acquisition. He starts with his Japanese object-verb VP word order and then he shifts to the correct English verb-object word order. These substages are labelled VPi and VPii since the only change in the learner's syntax is the word order of the VP. There is also evidence

9

from L2 learners of English who start with head-final/object-verb VP Turkish (Haznedar, 1997) and Farsi (Mobaraki, Vainikka & Young-Scholten, 2008). These utterances are similar to the examples from children learning English as their first language: there is no evidence for any functional syntax.

(6)   a. VPi: Japanese object-verb (OV) order

       bread eat

       bananas eating

     b. VPii: English verb-object (VO) order

       eating banana

       wash your hand

After the learner's initial reliance on their first language, they begin to acquire the inflectional morphology and syntax of the target language, in other words, they begin to add functional projections to their interlanguage English. The order in which they do so is common for learners of a given target language regardless of the learners' L1s, their age, the context of their exposure or their educational background (Hawkins, 2001). As with children learning their first language, internal linguistic mechanisms drive acquisition as the learner (subconsciously) responds to the input around him or her. Under the theory of Organic Grammar, the learner's consistent (but not necessarily always accurate) production of inflectional morphology represents the associated syntax. That is, the learner's consistent production of past-tense –ed indicates the learner's interlanguage has a TP and the learner's consistent production of 3rd person singular –s indicates an AgrP in his/her interlanguage system. The criteria for the acquisition of subject-verb agreement in English use of –s to consistently distinguish it in function from other morphemes, its use with more than one stem and in more than one linguistic context and its use in contexts which require it.

Several correct uses of a form by a learner do not indicate its acquisition; these could be use of unanalyzed forms by second language learners, for example Myles (2004) in her study of older children learning French in the classroom and Wagner-Gough (1978) in her longitudinal study of a young immigrant boy learning English. Depending on the language, there may also be syntactic indicators for the learners' acquisition of AgrP, e.g. nominative case marking for pronomination subjects.

The steps in the acquisition of English morphosyntax are shown in (7). Under Organic Grammar, each syntactic projection in the tree in (1) above is recognized as a stage of syntactic development, and these stages move from the bottom of the tree upwards. The first step for the learner is identifying in the input the head of a projection. With continued input, the learner's innate capacity for language, in the form of Universal Grammar, then provides the syntactic structure. We assume - along with many other adult second language acquisition researchers - access to Universal Grammar across the lifespan. For Organic Grammar, this means that apart from initial first language influence for VP word order, the same stages are predicted for children and adults, for both first and second language acquisition (for more evidence on the adult second language acquisition of English, see Young-Scholten and Ijuin, 2006).

(7)   Stages in the acquisition of English under Organic Grammar

VP-stage: verb and its arguments (bare verb, or a single form of verb occurs –

e.g. *is* or *are* for *be*); first L1 word order, followed by L2 word order

AspP-stage: the *–ing* suffix on the verb, contrasting with a bare form (no auxiliary yet);

non-nominative subjects as well as optional nominative subjects might occur up to the AgrP

stage

PerfP-stage: participle form of the verb (*-en;* without auxiliary) – possible aspectual

contrast with *–ing*

NegP-stage: Sentential *not* follows subject DP (or, *don't* overgeneralized to all tenses and persons)

TP-stage: past tense *–ed* acquired on regular verbs (also, auxiliaries are expected to emerge - beyond *don't* - but not always correctly marked for person agreement yet)

AgrP-stage: person agreement (in particular, 3rd person sg. *–s*) acquired; person forms of *be* acquired (*am, is, are*; *was, were*); null subjects no longer possible; non-nominative subject pronouns no longer possible

CP-stage: embedded clauses with an overt complementizer and full structure within the embedded clause; object wh-questions with full structure

Psycholinguistics and L2 processing

In the quest to solve the mystery of missing inflections, researchers usually collect oral and sometimes written production data from L2 learners. Data might have been spontaneously produced during conversation with the interviewer or during a range of specific tasks with learners' production recorded and then transcribed. The problem with such data is that they probably do not reveal the full extent of learners' underlying syntactic competence. Researchers have therefore increasingly turned to psycholinguistic, experimental techniques which examine on-line processing during listening or reading (Clahsen, 2007; Clahsen, Felser, Neubauer, Sato & Silva, 2010). There is now a range of options which measure responses to not only syntactic stimuli but a range of other linguistic stimuli (see chapters by Marinis & Cunnings, and by Wright in this volume). When it comes to past tense and subject-verb agreement, processing studies confirm L2 adults' difficulties. Studies by Chen, Shu, Liu, Zhao and Li (2007), McDonald (2000, 2006), McDonald and Roussel (2010) and Sato and Felser (2008) used different techniques and showed interesting yet sometimes contradictory results for adult second language learners from a variety of first language backgrounds. In general, these processing studies confirm what the oral and written production

studies show. However, the focus of these studies has not been on identifying the source of learners' variable production of forms. One way to make progress in solving the mystery is to collect not only production data but also perception and processing data from the same group of adult second language learners.

Solving the mystery: Methodology

The data presented here are from Kahoul's (2014) PhD for which L1 speakers of Arabic and Chinese at varying levels in their L2 English were compared. The two languages differ in important ways. Arabic marks both subject-verb agreement and past tense while Chinese marks neither. Under the Full Transfer/Full Access hypothesis, Arabic learners are predicted to transfer TP and AgrP to their L2 English. If there is variable production of past tense or agreement suffixes, this will be a case of Missing Surface Inflection rather than an absence of TP or AgrP. When it comes to Chinese learners, they cannot rely on their native language except perhaps for reassembling their Chinese aspect feature as tense; see above. This is because Chinese lacks TP and AgrP. Their continued access (=Full Access) to Universal Grammar will help them acquire these features. Thus Full Tranfer/Full Access predicts that Arabic learners will exhibit a different and well as a more rapid acquisition trajectory than the Chinese learners of English. L1-based variation will also be due to what researchers refer to as Prosodic Transfer: like English, Arabic allows clusters of consonants at the ends of words but Chinese does not. This is important because in attaching to stems, tense suffixes and agreement suffixes which are single consonants often attach to stems with one or more consonants as in 'She talked loudly on the bus' [lkt] and 'He bends during yoga' [ndz]

Like Full Transfer/Full Access, Organic Grammar predicts that Arabic and Chinese learners will acquire TP and AgrP but that they will follow the same route of acquisition. Thus while Full Transfer/Full Access predicts superiority for Arabic learners from the early stages onwards because L1 gives them TP and AgrP, Organic Grammar predicts parity for Arabic and Chinese.

Under Organic Grammar, acquisition of TP and AgrP in a second language is based solely on ability to acquire new syntactic structure from scratch. This is known as X'-Theory and is part of Universal Grammar. Learners do so in response to the English input to arrive at the tree in (1). Finally, deficit hypotheses, where adult second language learners are no longer able to access Universal Grammar and cannot therefore acquire new features such as tense and agreement, predict that Chinese learners will perform badly from the start and at all further levels of proficiency, and much worse than Arabic learners.

Kahoul (2014) recruited 34 speakers of Arabic (12 Syrian, 8 Libyan, 5 Iraqi, 4 Egyptian, 3 Jordanian and 2 Saudi) and 37 speakers of Chinese who were improving their English in the UK. He then placed them at three different levels of English proficiency in a cross-sectional study meant to mimic actual development over time. At the time Kahoul collected data, they had resided in the UK from two months to eight years and had started learning English in school in their home countries after the0 age of seven. Thus they were all instructed learners who had then received at least some if not considerable amount of input in the target language country. In a cross-sectional study, the researcher needs to make sure we use as robust as possible an objective test of placing learners at distinct proficiency levels, if the researcher intends to show how actual acquisition over time occurs. Researchers rely on various oral and written tests such as the Oxford Quick Placement Test used to place students on English-as-a-second-language programmes. However, Kahoul adopted the more rigorous test developed by Unsworth (2005, 2008) as a three-step assessment resulting in an Age-Sensitive Composite Proficiency Score. Table 1 below shows Kahoul's placement of the 71 learners at what is standard in research: three proficiency levels, low (beginner) mid (intermediate) and high (advanced):

Table 1 Proficiency groups

	Arabic speakers	Chinese speakers
Low	11	13
Mid	14	13
High	9	11

Kahoul posed two research questions: (1) Is learners' oral production of tense and agreement consistent? (2) Do learners' data from perception and processing tasks reveal the same patterns as the oral production data? The tasks Kahoul gave his participants were elicited imitation and picture-choice tasks which were supplemented with a measurement of reaction time and eye-tracking.

Reaction Time measures the individual's speed in responding to test stimuli which indicates processing load. Mastery of a linguistic feature – a feature that comes to be represented in the learner's mind – results in rapid, automatic performance while non-mastery involves slow responses. The researcher's decision on whether reaction times are rapid or slow is not absolute. Rather, times are relative and based on native speakers' performance on the same task. The tracking of an individual's eye movements in response to various stimuli also reveals information about what the viewer's mind represents. This is based on the observation that when we are presented with visual stimuli, we move our eyes to focus attention on what interests us or what is relevant. In addition to reaction times, eye movements provide the researcher with insights into underlying cognitive processes (Hayhoe & Ballard, 2005; Liversedge & Findlay, 2000). The researcher can track participants' saccades (eye movement from one fixation or gaze to another) in response to visual stimuli which is presented with auditory stimuli to provide information about

real-time language processing with regard to a range of specific linguistic features (Tanenhaus & Trueswell, 2006). For example Arnold *et al.* (2000) looked at how gender information inherent in English pronouns influenced moment-by-moment interpretation. Participants saw pictures containing choices between characters and heard pronouns referring to just one of them. Results showed that participants looked at the target character in the picture within 200 milliseconds after they heard the pronoun and their eyes fixated longer on the target character than on the incorrect one, the so-called competitor. In the present study, Kahoul undertook these two eye movement measures: first look and length of look.

Kahoul tested oral production with an elicited imitation task where participants heard 50 recorded sentences (mean 12 syllables in length) and had to repeat them. To avoid rote imitation and to check participants' comprehension of the sentence, the participant then saw three pictures and had to choose the correct picture before repeating the sentence (see Erlam, 2006, 2009; also Marinis & Cunnings, chapter 10 for more on the elicited imitation technique). The following is an example of a sentence imitation + picture item:

(8) Yesterday, John ate his breakfast before he went out.

Results

We address each research question: (1) Is learners' oral production of tense and agreement consistent? (2) Do data from perception and processing tasks reveal the same patterns as the oral

16

production data? In the process of addressing these questions, we also consider the three hypotheses discussed above, Organic Grammar, Full Transfer/Full Access and deficit hypotheses.

Oral production

Table 2 below shows the results for correct past tense production and Table 3 below shows the results for subject-verb agreement. For past tense, comparisons between the same-proficiency different-first language groups showed no significant language-based differences between the two groups at the two lower levels but there was a significant language-based difference between the two High groups (Arabic High vs. Chinese High)

Table 2 Correct production of past tense by Arabic and Chinese learners

	Arabic speakers				Chinese speakers		
	Score	%	SD		Score	%	SD
Low (n=11)	177/317	55.83	24	Low (n=13)	193/411	46.95	16
Mid (n=14)	276/416	66.34	21	Mid (n=13)	202/399	50.62	22
High (n=9)	275/297	92.59	5	High (n=11)	269/362	74.30	13

For subject-verb agreement, for 3rd person singular –s, the Low and Mid groups did not differ significantly based on their native language but at the highest level of proficiency, the Arabic learners of English were significantly better at producing agreement marking than their Chinese counterparts.

Table 3 Correct production of verbal agreement by Arabic and Chinese learners of English

Arabic Speakers				Chinese speakers			
	Score	%	SD		Score	%	SD
Low (n=11)	140/298	46.97	27	Low (n=13)	192/354	54.23	11
Mid (n=14)	170/342	49.70	20	Mid (n=13)	168/352	47.72	23
High (n=9)	226/260	86.92	8	High (n=11)	217/298	72.81	12

These oral production task results support Organic Grammar for the Low and Mid learners: at the lower levels of proficiency, learners have not acquired TP and AgrP. The Full Transfer/Full Access hypothesis is rejected as it predicts differences at the early stages under the assumption that only Arabic speakers can transfer their TP and AgrP. A comparison of percentages of correct suppliance of tense and of agreement forms by learners at the Mid level points to the acquisition of TP preceding the acquisition of AgrP, as predicted by OG, as per the tree in (1) and the steps in (7). There is an obvious native language effect at the highest proficiency levels: Chinese speakers lag behind their Arabic counterparts. There could be phonological reasons for this; recall that Chinese does not allow word-final consonant clusters. The Prosodic Transfer Hypothesis was mentioned above in connection with the Missing Surface Inflection Hypothesis and Full Transfer/Full Access, but it is also compatible with Organic Grammar. Chinese learners' oral production of forms ending with consonant clusters is predicted to be worse than Arabic speakers' oral production of these, and forms without consonant clusters are predicted to be more successfully produced by learners. Kahoul's analysis of the elicited imitation data thus looked at the production of suffixes by **allomorph** and by regular and irregular verbs (shown in Table 4 below).

No significant phonological effects were observed in the results of the Chinese speakers, but the Arabic speakers displayed some phonological effects in their production of regular past tense:

allomorph type was significant for the consonant + t allomorph, significantly lower than for consonant + d allomorph at the Low and Mid levels, and the consonant + t allomorph was significantly lower than the CVV-d allomorph at the Mid level. There were no significant differences for any other combinations. Production of consonant cluster vs. non-cluster forms showed no significance at the Low or High levels, but did at the Mid level. A comparison between participants' production rates of past on regular and irregular verbs also revealed no significant differences at any of the three levels for the Arabic speakers. However, the Chinese speakers showed significantly higher target-like production of irregular verbs than regular verbs at Low and Mid levels, but not at the High level. If native language phonological constraints act on production of clusters in verbs with regular –ed, rates for irregular verbs should have been higher. Therefore, the advantage of Arabic learners over their Chinese counterparts is only for High-level learners where the Chinese High-level learners' production of regular vs. irregular morphology no longer differs statistically from the Arabic learners'.

For 3rd person singular –s agreement, rates of production for Arabic proficiency groups with allomorph type as a factor showed no significant differences at any level. For the Chinese learners, at the High level, one of the allomorphs, a short vowel followed by –z, showed significantly better production than the CVV-z allomorph.

Table 4 Correct production of regular vs. irregular past tense verb morphology suffixes

	Regular			Irregular		
Arabic	Score	%	SD	Score	%	SD
Low (n=11)	111/202	54.95	27	66/115	57.39	23
Mid (n=14)	183/278	65.82	24	93/138	67.39	17
High (n=9)	193/211	91.46	7	82/86	95.34	6
Chinese						
Low (n=13)	118/275	42.90	19	75/136	55.14	15
Mid (n=13)	119/262	45.41	26	83/137	60.58	16
High (n=11)	177/248	71.37	16	92/114	80.70	9

Perception

     To compare participants' oral production data with perception and processing data relating to tense and agreement, a computerised picture-choice task presented individual learners with arrays of three pictures related to various times of day (for tense) and various numbers of people (for agreement). There was a target picture (=the correct choice), competitor picture (= a possible but wrong choice) and a foil (a completely unrelated picture to make sure the participants were following instructions). Participants listened to sentences matching the target and for each sentence, they had to indicate which picture matched. A distractor item or trial which was not included in the data analysis first introduced the context of time; the participant's response to the 9

o'clock picture indicated the participant's understanding of the tense context.
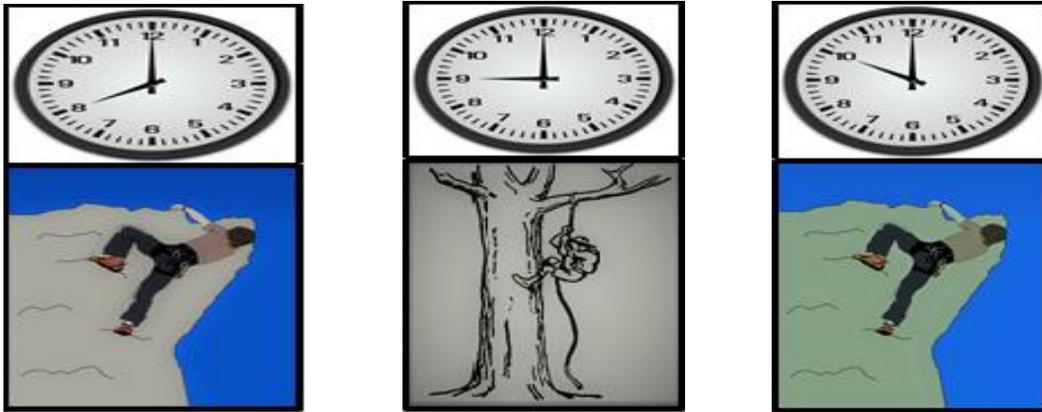
(9)   a. Distractor Trial

It's ten o'clock right now. Listen and choose:

"He was climbing something one hour ago."

b. Experimental Trial

"He climbed up the mountain."

Establishment of the time as 10 o'clock and perception of the past tense inflection on the verb 'climb' in (b) was expected to prompt participants to choose the 8 o'clock picture.

Subject-verb agreement was tested with one oral stimulus, two stimuli depicting the same action and (dancing in the example below) one foil picture (playing football in the example below). To prevent the participants from using the subject as a clue, the subject of each sentence was not provided but rather masked with silence.

(10) "Every party [….] dances happily"

Perception of the –*s* agreement inflection on 'dances' was expected to prompt participants to choose the target picture, the single character who dances at every party. The task consisted of 88 items or trials with 10 training trials, 27 experimental trials, 39 distractors and 12 experimental-like fillers. A high number of distractors and fillers is necessary in such tests to prevent participants from figuring out what is being tested and monitoring their performance.[vii] The sentences in the experimental trials contained verbs inflected for either agreement or past tense. The task contained 15 items inflected for past and 12 items inflected for agreement. As in the production task, allomorphs of these suffixes were also manipulated to see whether the phonological form of the suffix had an influence. Unlike the production task, a control group of ten similarly-aged native English speakers was included to confirm test validity and to compare the second language learner groups' behavior with that of native speakers. The results show that all three Chinese groups and the Low and Mid Arabic groups had problems perceiving past tense, and all apart from the Arabic High group were significantly different from the native speaker group.

Table 5 Successful perception of past tense by Arabic and Chinese learners

Arabic				Chinese			Native		
	Score	%	SD		Score	%	SD	Score	%	SD
Low (n=9)	77/135	57.03	19	Low (n=11)	102/161	63.35	20	139/150	92.66	10
Mid (n=13)	133/195	68.20	26	Mid (n=13)	107/180	59.44	24			
High (n=9)	117/135	86.66	10	High (n=10)	104/150	69.33	26			

Perception rates for past tense for separate Arabic and Chinese speaking groups by proficiency sub-group showed a significant difference among the Arabic Low and Mid vs. the Arabic High group. Perception of agreement was similar: there was a significant difference among the two lowest Arabic groups and the highest group, and this group (unlike all other groups) also did not significantly differ from the group of native speakers. For the Chinese groups, no such effects were found in the perception of past or agreement.

Table 6 Successful agreement perception by Arabic, Chinese and native speakers

Arabic				Chinese			Native		
	Score	%	SD		Score	%	SD	Score	%	SD
Low (n=9)	80/108	74.07	16	Low (n=11)	104/130	80	10	117/120	97.5	4
Mid (n=13)	112/156	71.79	18	Mid (n=13)	118/156	75.64	15			
High (n=9)	100/108	92.59	9	High (n=10)	94/120	78.33	18			

Native language prosodic constraints might also be acting on perception. As noted above, the

perception results were also examined by suffix allomorph, cluster vs. non-cluster and regular vs. irregular. These variants revealed no significant differences for the Arabic speakers; that is, phonology had no influence. For the Chinese, there were significant differences, but only for Low level, where perception of a long vowel preceding the consonant -d marking past tense was significantly lower than that of a short vowel CV-d and for Mid-level, perception of a consonant followed by the consonant -t was significantly higher than for the other allomorphs.

In the perception of third-person singular –s, there were no significant differences for Arabic learners at any level. The Chinese results also showed no significant differences for cluster vs. non-cluster at any level, and no significant differences for Mid and High level learners with respect to allomorphs. However, for Low level Chinese learners, the consonant –z following a long vowel was significantly better than other allomorphs and the consonant –z following a short vowel was significantly worse in perception than the C-z allomorph.

Processing: Reaction Times

The picture-choice task also measured learners' reaction times for further insights into their underlying linguistic competence. Only the reaction times on the experimental trials which were answered correctly are of interest and are therefore presented here. Similar to the results discussed thus far, this measure did not show any differences between the two native language groups by proficiency level. For each group, results demonstrate learner improvement on past tense with rising proficiency, but not to native-speaker level. For the Arabic learners, there were no significant differences between Low- and Mid-level and Mid- and High-level, but there were significant differences between Low- and High-level groups. Chinese results showed a significant difference among proficiency where Low- and Mid-level learners performed similarly as did Mid- and High-level learners, and High-level learners performed significantly better than Low-level learners. For agreement, development of English as measured by faster reaction times by proficiency level revealed significant differences between all proficiency groups and the native speakers, no

24

significant difference between the Low- and Mid-level groups but reaction times showed that the High-level groups performed significantly better than the Low-level groups

Processing: Eye Tracking

This measure provided the most interesting data in the entire study in terms of native language and proficiency effects. Length of eye-look and first look to target measure were applied to the perception data described above. Length of eye-look, as shown in milliseconds in the figures below, shows to what extent learners prefer the target picture over the competitor (eye looks to the foil is unrelated to tense or agreement). First look is a measure of how rapidly learners spot the target picture after hearing the sentence.

For past tense, the Low-level Arabic and Chinese speakers looked for equal lengths of time at target and competitor, and Mid- and High-level Arabic speakers – but not Chinese speakers – spent a longer time looking at targets. They patterned with native English speakers who looked significantly longer at targets than competitors.
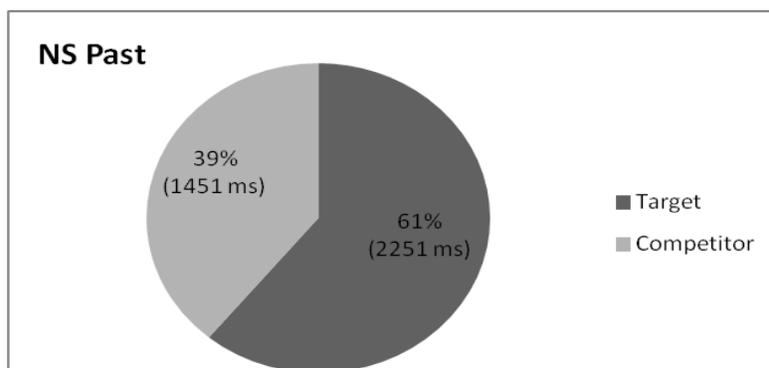
Figure 1 Eye-looks in past tense items by native English speakers

The Arabic and Chinese Low-level groups initiated looks to both target and competitor pictures in the same manner, with no apparent preference for either, as shown in Figure 2.
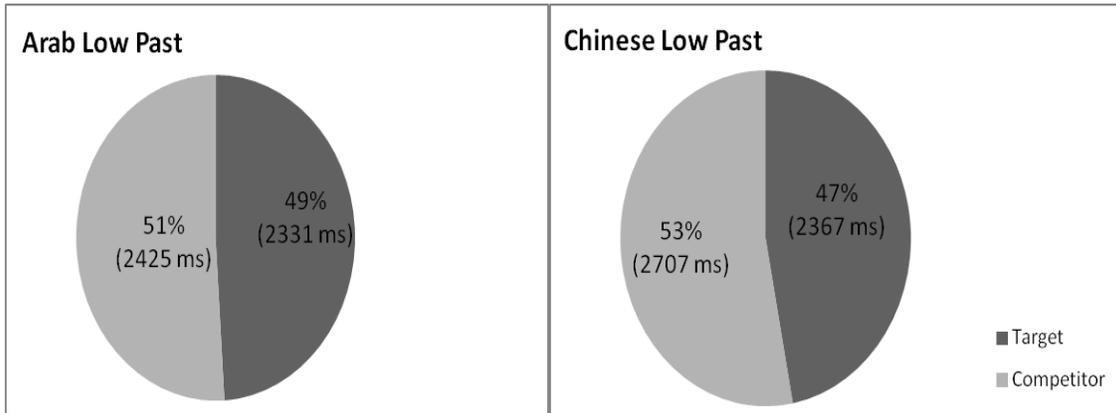
Figure 2 Eye-looks in past tense items by Low-level non-native groups

Mid-level Arabic learners looked significantly more at targets than competitors while there were no significant differences for Chinese learners: they looked at target and competitor pictures in the same manner.
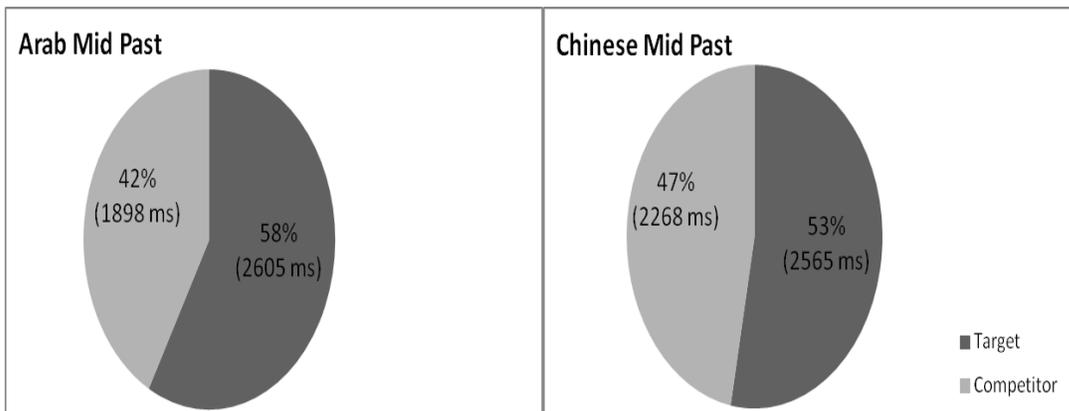

Figure 3 Eye-looks in past tense items by Mid-level non-native groups

Figure 4 shows that Arabic and Chinese High-level learners patterned differently and statistical tests show that, while the Arabic group looked at targets significantly more than at competitor pictures, the Chinese group did not.
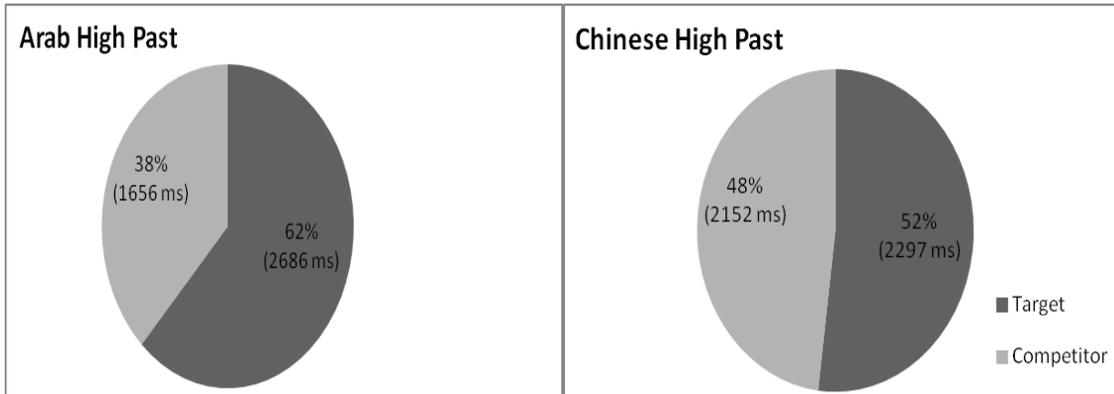
Figure 4 Eye-looks in past tense items by High-level non-native groups

While past tense showed no preference for either native language group at the lowest proficiency level, and at the two higher levels only for the Arabic group, the eye tracking of subject-verb agreement showed significant preference for the target picture by all groups. For the native English speakers and for all three proficiency levels and for each language group, eye-looks at target pictures were significantly more frequent than at competitor pictures.
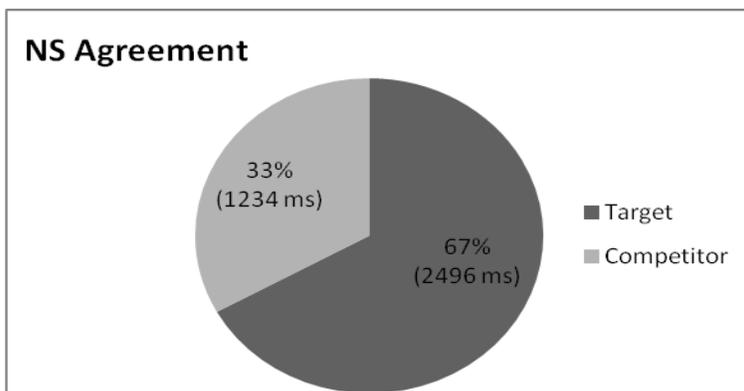
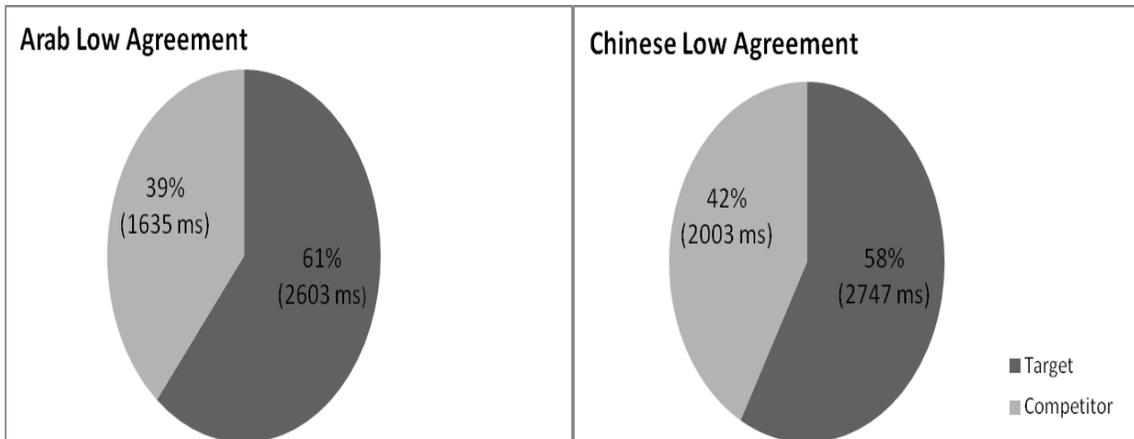**Figure 5** Eye looks for agreement by native English speakers

Figure 6 Eye-looks in agreement items by Low-level non-native groups
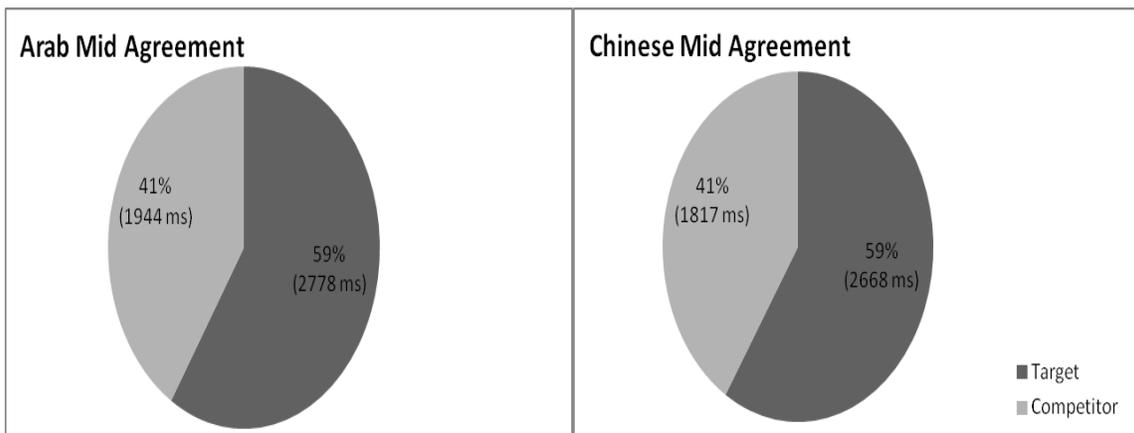

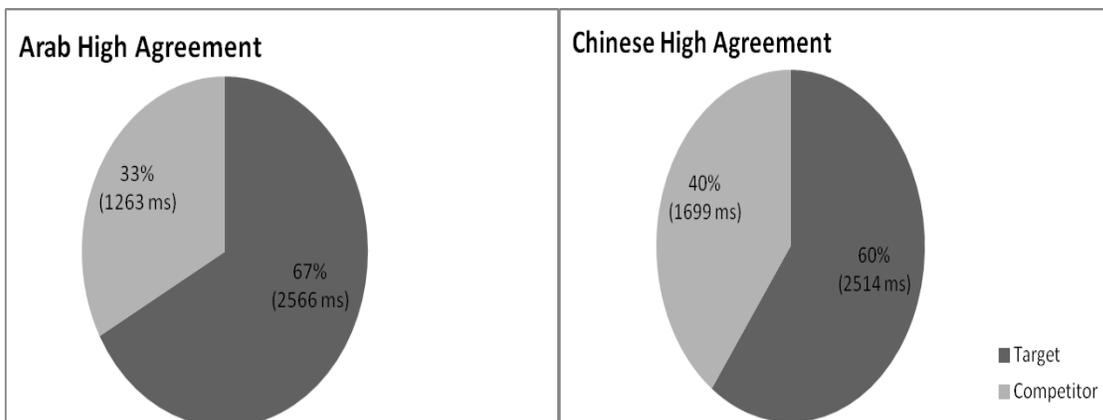**Figure 7** Eye-looks in agreement items by Mid-level non-native groups

Figure 8 Eye-looks in agreement items by High-level non-native groups

In the measurement of first look, i.e. how quickly a participant looks at the target picture immediately after presentation of the stimulus, Arabic speakers' sensitivity in past tense trials, as shown by speed, increases with proficiency: there were no significant differences between the Low- and Mid-level groups but the High-level group differed significantly from the Low-level and Mid-level groups. The Chinese speakers' speed for first look increases significantly from Low to Mid levels, but then decreases for the High-level learners. Results for native language showed significant differences between the Arabic and Chinese groups at Low level and High level but not at Mid-level. Results for the native English speakers vs. non-native speakers showed no difference between the native speakers and the Arabic High-level group but significant differences between the native speakers group and the other non-native groups.

Arabic and Chinese learners' speed with respect to agreement increased with rising proficiency; however at the highest proficiency level, Chinese learners' sensitivity then decreased. Similar proficiency groups in each language were not significantly different at either the Low-level or Mid-level but at High-level, Arabic learners' sensitivity to agreement was significantly higher than that of their Chinese counterparts and the Arabic High-level learners and native speakers hardly differed with their scores approaching significance.

Discussion

Considering the clues to solving the mystery

What do these results tell us about missing inflections? By all measures discussed above, Low- and Mid-level learners of two very different native language backgrounds followed a similar route in acquisition of TP and AgrP, as predicted by Organic Grammar. No significant differences were found between the two groups at these proficiency levels, neither in production nor in the various measures of perception and processing. A significant difference emerged at the highest

level of acquisition, however. The Low-level Arabic and Chinese speakers produced agreement and tense morphology variably, they experienced perceptual limitations, they exhibited slower reaction times, slower first looks and they showed no gaze preference for the target pictures during eye tracking. At the Mid-level, they all improved and the same patterns were observed.

With increasing overall proficiency, the Arabic and Chinese speakers began to differ: at High level, Chinese speakers continued their variable oral production while Arabic speakers became consistent. All speakers' perception showed improvements at the highest level. No differences were found for reaction times, but for eye movement, patterns for past tense for Arabic speakers showed a significant preference for the target and significantly faster first looks not only for past tense but also for agreement. Table 7 shows the group mean results underlying the results of the tests for statistical significance referred to above, namely the percentages for accuracy for production and perception, correct eye-looks at target and the milliseconds in reacting to the picture stimulus and in looking first at the correct target picture.

Table 7 Development of TP and AgrP (group means)

		Arabic Low	Arabic Mid	Arabic High	Chinese Low	Chinese Mid	Chinese High	Natives
Production (%)	Past	56%	66%	93%	47%	51%	74%	-
	Agr	47%	50%	87%	54%	48%	73%	-
Perception (%)	Past	57%	68%	87%	63%	59%	69%	93%
	Agr	74%	72%	93%	80%	76%	78%	97%
Reaction Time (ms)	Past	2668	2153	1503	2729	2640	1854	1017
	Agr	2485	2300	1379	2468	1867	1541	857

Looks at target (%)	Past	49%	58%	62%	47%	53%	52%	61%
	Agr	61%	59%	67%	58%	59%	60%	67%
First look at target (ms)	Past	1246	940	773	1654	1126	1425	695
	Agr	1109	933	741	1195	878	1046	589

Solving the mystery?

Variable production of third-person singular –s and past tense morphology by the Chinese learners at the Low, Mid and High proficiency levels and by Arabic learners at Low- and Mid-level, coupled with perceptual limitations of the same morphological items, indicate that the source of learners' problems is representational, i.e. due to learners' morphosyntax lacking a TP or an AgrP. Learners' development with rising proficiency, particularly for Arabic-speaking learners in their oral production and perception of morphology, indicates a gradual building of syntactic representations. That is, the Low- and Mid-level Arabic speakers and all levels of Chinese speakers gradually acquire TP and AgrP. The Arabic speakers show no initial or intermediate advantage over the Chinese speakers but pull ahead at High level. This difference cannot be traced to phonology in any obvious way. The data presented here thus do not provide support for the Missing Surface Inflection Hypothesis. The comprehension and processing experiments would have revealed competence in these suffixes, contrary to what was found: production, comprehension, reaction time and eye tracking provide consistent results at the various proficiency levels. Deficit approaches, and Full Transfer/Full Access (whereby all functional projections are present in the L2 learner's internal grammar throughout acquisition) are thus not supported by these data. Full Transfer is not supported: Arabic learners were expected to but did not produce, perceive and process third person singular and past tense morphology better than their Chinese counterparts

from early on. They only did so at High level. The data from the Low- and Mid-level Arabic and Chinese learners of English support Organic Grammar in that absence of TP and/or AgrP explains these patterns whereas the High-level Arabic learners' patterns point to their acquisition of TP and AgrP.

No hypothesis discussed thus far predicted what we found for the High-level learners. There are two possibilities here: (1) Hawkins' (2001) Modulated Structure Building, according to which learners build structure and when functional projections are acquired during development, the learner's native language can provide assistance in acquiring a new projection; or (2) the High-level Chinese speakers, in fact, are high intermediates, and there are speakers at a more advanced level who acquire TP and AgrP. In Hawkins' version of structure building, it is not clear exactly what the predictions would be for Chinese learners of English, since the Arabic and Chinese learners pattern the same at lower levels.

For Chinese High-level learners, it is possible there is a stage at which TP and/or AgrP is being acquired, although later than the Arabic learners: the U-shaped learning curve shown in the 'Sensitivity to Past' graph in Figure 9 for the picture choice task results mirror the tense/agreement of children's first language acquisition of French reported in Legendre and Davidson (2003). They argue that TP is acquired first, then once agreement shows up, there is competition for space (no AgrP yet) and performance on tense declines before TP and AgrP are then both projected.
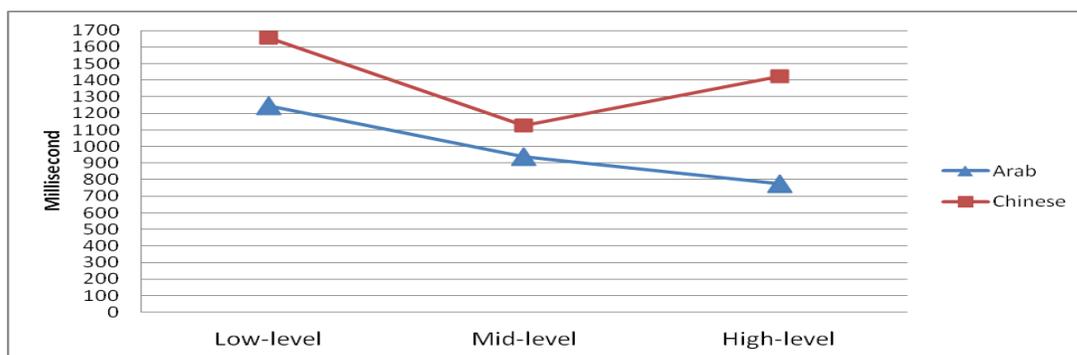

Figure 9 Learners' perceptual sensitivity to past tense (picture choice task)

The production results indicate that for the Arabic learners, TP is acquired before AgrP; for Chinese learners, AgrP is not acquired until TP is. However, the perception and processing results do not seem to support this. One possible reason for this discrepancy is the role of the perceptual salience of the sibilant fricatives [s] and [z] in facilitating processing of third person singular agreement as in the four allomorphs in 'sits', 'loves', 'crushes' and 'knows'. Therefore, (i) the oral production data present evidence for the acquisition of TP before AgrP, but (ii) the perception and processing data cannot be considered evidence against the acquisition of TP before AgrP because the perceptual salience of the agreement suffix allomorphs makes the comparison unreliable. We also have to take into consideration that all of the learners in this study were instructed prior to moving to the UK, and instruction in agreement and past tense is inevitable.

Our results do not warrant the conclusion that adult L2 learners have only partial access to Universal Grammar, i.e. they do not support the deficit hypotheses discussed above. At the early stages of development, both language groups pattern extremely similarly. Under Organic Grammar, with direct access to Universal Grammar, this is predicted. Why exactly Chinese speakers' development initially proceeds like Arabic speakers' development but then stalls remains another mystery to be solved. We suggest this might be investigated by looking closely at the minor phonological differences detected in the oral production and perception tasks. These might reflect learners' phonological problems in processing the target language input as they search for information in the input to trigger their acquisition of TP and AgrP in English (for elaboration on this idea, see Vainikka & Young-Scholten, 1998). In this sense, too, L2 learners are detectives in solving their own mystery of missing inflections.

---

[i] Much recent Minimalist theorising (e.g. Chomsky 2008) excludes AgrP for theory-internal reasons, but we continue to assume AgrP.

[ii] An anonymous reviewer points out that utterances such as (2) and (3) might reflect fuller adult sentences the child hears, e.g. "Do you want to give doggie a bone?" without missing inflection.

[iii] Referred to as Root Infinitives (Rizzi, 1993/4); Optional Infinitives (Wexler, 1994); Root Defaults (Paradis & Crago, 2000).

[iv] Further examples of non-adult syntax (from Hyams, 2007; Gagarina & Gülzow, 2008; Castro & Gavruseva 2003) include

    (i) *Papa schoenen wassen*.　　　　　　[Dutch]
      Daddy shoes　wash-INF
    (ii) *Michel dormer*.　　　　　　　　　[French]
      Michel sleep-INF
    (iii) *Jag ocksa hoppa　　där　a　där*.　　[Swedish]
      I　also　hop-INF　there and there
    (iv) *lomat　　koleso*.　　　　　　　[Russian]
      break-INF wheel-ACC
    (v) *Mi hacer　　otra*.　　　　　　　[Spanish]
      me make-INF another

[v] This is Rizzi's (1993/4) Truncation Hypothesis and Wexler's (1994) Optional Infinitive Stage; see also Wexler, Schütze and Rice (1998)

[vi] Child Language Data Exchange System (MacWhinney & Snow 1985; http://www.childes.psy.cmu.edu) is an open-source databank of orthographically-transcribed, coded speech samples from children acquiring a range of languages. Anyone can query it for counts or co-occurrences for syntax, morphology and lexis.

[vii] Six were similar to the past tense trials and six were similar to the verbal agreement trials. These were experimental-like in that they used the same auditory stimuli of the experimental trials but with bare verbs.
The rationale for this was to detect if participants considered all verbs presented inflected and gave responses based on this, which would be a deceptive indicator of their perception. The inclusion of these items meant that they did not so do, i.e. this strategy was successful.