Wh-questions: Moving beyond the first phase

Jill G. de Villiersa,*, Peter A. de Villiersa,1, Thomas Roeperb,2

Psychology Department, Smith College, Northampton, MA 01063, USA
Linguistics Department, South College, University of Massachusetts, Amherst, MA 01030, USA

1. Introduction

In this paper we put forward a theory that ties together several strands in the acquisition of wh-movement from complex sentences. First, it makes use of a feature, Point of View (PoV), that was proposed to accommodate some phenomena about children’s mastery of complementation. Second, the account differentiates the developments that allow long distance movement and indirect questions. We will build a Feature Bundle to specify exactly how PoV interacts with other features to predict both grammatical facts and the acquisition path. Third, we predict serious delay in children with specific language impairment (SLI) on the development of this intricate system. Fourth, this mechanism enables us to make a specific prediction about a dialect difference between Mainstream American English (MAE) and African American English (AAE) that is borne out by new empirical data from a large sample of children speaking each dialect.

Our expository method has several goals: we will provide an intuitive account that does not depend upon a specific notation, then we develop a Feature notation that captures the mechanisms involved in the syntax and acquisition stages, which should eventually connect to the semantics interface.

First, consider the properties surrounding PoV that are distinctive to complement structures under communication and mental verbs. de Villiers (2001, 2005) argued that a complement under certain mental or communication verbs takes a separate PoV in its CP, that is, the truth of the lower clause is relative to the subject of the matrix clause, i.e. Martha in:

The paper presents a feature-checking theory of wh-movement that attempts to accommodate both adult grammar and the path of acquisition by which children handle long distance movement, indirect questions and partial movement. Partial movement is not a grammatical option in English but it is adopted as an option in development. The account makes several predictions about the performance of children with specific language impairment (SLI), and also predicts a particular advantage for children who speak African American English (AAE) over those who speak Mainstream American English (MAE). The empirical data are taken from a study of 590 children, both typically developing and language-impaired, and both AAE and MAE speaking, aged four to nine years. The tasks involved answering wh-questions after stories as part of the field-testing of a new language assessment instrument. The questions included multi-clause questions with or without medial wh-complementizers. The predictions are borne out that children with language impairment have prolonged difficulty with real long distance movement and medial questions, and that children who speak AAE are at an advantage in avoiding certain errors (partial movement) because of the dialect’s characteristic marking of indirect questions via inversion in the lower clause.

© 2010 Elsevier B.V. All rights reserved.
1) Martha thought [her dog was a genius].

If a wh-question is asked, and the wh appears in the top CP:

2) What did Martha think her dog was?

the only legitimate answer is a genius, whatever the speaker might believe about the dog in question. The wh-question has inherited the PoV from the lower clause, namely Martha’s PoV, creating a classic “opaque” utterance, one that will not tolerate referential substitution, for example (see de Villiers, 2001, 2005; Hollebrandse and Roeper, 1998; for extensive relevant discussion).

In several papers, Jill de Villiers has made the claim that a developing child of age three years or so lacks this PoV feature on the lower clause under verbs such as think and say, and as a result of this incomplete representation, the child is not able to understand questions such as 2). For instance, asked:

3) What did the mother say she bought?

three year olds mistakenly answer what she actually bought, not what she said she bought. There is then a further consequence, that children who lack this linguistic representation also fail tests of false belief, or a mature Theory of Mind, arguably because they lack the linguistic means to represent another’s false beliefs for purposes of explicit reasoning (de Villiers, 2005, 2007; Roeper, 2007).

A second well-established fact about acquisition eludes simple integration into the same account, namely, how do children interpret medial wh-words as in (4), which do not correspond to a real question:

4) How did she ask what to bake?

A large range of work has shown that children provide what is called a partial movement analysis, answering what and treating how as a wh-scope-marker (the “medial” answer) as found in many languages (Fanselow, 2005; McDaniel, 1989). This medial error persists longer in children than the error on opaque sentences as in 3) above that only have a trace in the medial position (de Villiers and Pyers, 2002) and is found across modalities of production (Thornton, 1991) and judgment (McDaniel et al., 1995) and in other non-partial movement languages such as French (Oiry and Demirdache, 2006). The syntactic basis of our earlier account (de Villiers et al., 1990; Roeper and de Villiers, 1994) rested on parametric theory: children choose “partial movement”, a Germanic option, rather than full movement, which is the English option (McDaniel et al., 1995). This account, though included in our current view, does not provide the kind of deep conceptual motivation that we think acquisition evidence deserves, nor did it integrate the PoV-properties of opacity into the system. In the present account we attempt an integration of these two phenomena of language acquisition and provide a first attempt at a more formal analysis.

The account of the errors needs to be integrated into a formal account of the subordination system whereby features from a higher verb are projected onto the CP of the lower verb, which in turn govern what forms of movement are allowed, and into the system of Phases and Feature Satisfaction. In modern accounts of syntactic movement, movement is motivated to satisfy a set of features that get checked off by moving elements into the position sharing that feature (Chomsky, 1995). The incompleteness of current accounts has motivated us to propose a more articulated account to handle both adult and child grammar and to accommodate the acquisition facts about the process of change.

The theory of cyclic long-distance wh-movement (Chomsky, 1977) requires what in 2) to pass through an intermediate CP, through the medial CP before proceeding to the upper CP, leaving a trace in that position. At some point in the derivation, the sentence was:

5) [Martha thought [what her dog was t]]

then it becomes:

6) What did Martha think [Spec-CP t2 [her dog was t1]]

The sentence 5) can never occur in surface structure. But other verbs do allow medial wh-words. Therefore we also need a mechanism to handle indirect questions under certain verbs:

7) The mother asked what to bake.

For inspiration we turn to the work of Chomsky (2005), which offers a theory of local transfer which converges with the acquisition evidence, though our hypothesis requires a small extension of the Transfer hypothesis. Chomsky argues that Transfer to the interface occurs for both phonology and semantics from the syntax at each Phase level, where Phase here is roughly a clause (a CP):

“there are Transfer operations: one hands the Syntactic Object (SO) already constructed to the phonological component, which maps it to the Sensori-Motor interface (“Spell-Out”); the other hands SO to the semantic component, which maps it to the Conceptual-Intentional interface. Call these SOs phases. Thus the Strong Minimalist Thesis entails that computation
We take the Transfer idealization to be that as much transfer to Interpretation as possible should occur at this point, namely at the first Phase. English and other long distance wh-movement languages represent, in a rough sense, a departure from the idealization of maximizing Transfer one Phase at a time. The theory of cyclic movement moves a wh-element to the Edge of the Phase, which forces the wh-word to be a part of the next higher Phase and therefore to receive its interpretation later. The Edge feature is what used to be called “the escape hatch” of Spec of CP. Commenting on a sentence from our experimental work:

3) What did the mother say $t_2$ she bought $t_1$

Chomsky (personal communication) says:

“In the example the trace $t_2$ is at the edge of the CP phase, so is not transferred at the CP level. CI therefore interprets only she bought what, where what is the copy left behind by movement, which receives a theta role and is available for reconstruction and for interpretation as a variable at the final stage. Note that the total semantic and phonetic interpretation has to be global; both CI and SM (say, prosody) enter into full semantic interpretation, and in both cases the whole expression has to be surveyed. Yes-no question intonation, for example, is only determined at the very end of the computation, long past the point where the intonation appears.”

Therefore $t_2$ is not completely interpreted in adult English because it is marked by the higher verb to move further. For the young child, in contrast, the wh is interpreted with respect to the lower verb only.

In what follows we argue that a subject-oriented PoV is projected as a part of a Feature Bundle from the matrix verb, e.g. think or say, to the lower clause, which forces a deferral of Transfer of the Edge feature to the next Phase. Our analysis fits other proposals (de Villiers, 2001, 2005; Hollebrandse, 1999, 2007; Hollebrandse and Roeper, 1998; Speas, 2000; Speas and Tenny, 2003) that the PoV element can be a single syntactic feature, as well as an element in a semantic representation. The Feature Bundle projected from the higher verb onto the lower CP determines the nature of the subordinate clause. For example, is it a direct question or an indirect question? Does it maintain the speaker’s PoV or become opaque? Without that feature projection, we argue that the Edge Feature will receive a local interpretation within the Phase. That involves interpreting the wh with respect only to the lower clause, that is, the interpretation is closed off or completed, and no longer subject to modification of its internal meaning. By default, it will be “true”, at least from the speaker’s perspective. At what point does the Edge Feature receive interpretation in the adult grammar? Chomsky’s comment raises the question of when the “whole phrase” interpretation occurs, which then requires interpretation of the Edge Features. In long utterances, where there is a series of adjuncts, it is commonly argued in psycholinguistics work that they are “shipped off” to semantic memory at clause boundaries. In computational terms, the speaker (perhaps particularly the child) seeks to Transfer the syntactic information to the Cognitive Interface as rapidly as possible. Therefore we argue that it is plausible that the child could regard the complement clause as one where the Edge Feature does undergo interpretation. We propose that Edge Feature interpretation can occur when 1) there is an explicit wh-expression AND 2) when the feature projection from the higher verb is not present.

Our hypotheses are twofold:

i. Full Transfer at the CI interface entails closing off the interpretation, a meaning we tentatively call ‘default factivity’. ii. Overt wh-expressions trigger interpretation of Edge Features in Spec of CP. If no feature projection from a higher verb has been recognized, then default factivity results automatically. In effect, then the interpretive apparatus for “the whole expression” can be locally invoked and the clause interpreted independently.

We now explore the consequences of this prediction in several environments. There are three environments to consider:

8) John wondered what Bill did.

b. Closed propositions:

9) John knew what Bill did.

Partial movement (like German):

10) What did Hans say who came?

Was hat Hans gesagt wer gekommen ist?

While the form in (c) is not English, partial wh-movement is common in many languages (see Fanselow, 2003 for an overview; also Abdulkarim, 2001; Dayal, 2000; Lahiri, 2002a, 2002b; Oiry and Demirdache, 2006; Schulz, 2004; for particular proposals). Chomsky (2005) proposes that it is the default form of cyclic movement not only for wh- but also for Tense, following Koster (2003).

We will now review these three constructions beginning with the contrast between indirect questions and closed propositions, which is a function of the higher verb. We cannot give a complete differentiation of the feature content of
projections from the higher V to the lower CP and its semantic consequences, but will just give those that play a role in our argument. Some facts need to be introduced to see the range of differentiation, but also, to see the challenge the child faces.

The first observation stresses that subordinate wh-complements are not necessarily indirect questions. If we compare:

11) John wondered where Bill went.
12) John knew where Bill went.

11) and 12) have a very different force. In 12), the lower clause refers to something that has been completed, *where Bill went* has a distinct reference, unlike the indirect question in 11), that is in contrast open or indeterminate. In the case of *know*, the embedded form does not have question force. Instead, the clause has been completed, or closed off in interpretation. There is a significant literature on this semantic point (for review see Hagstrom, 2003; also Égré, 2005; Groenendijk and Stokhof, 1982; Karttunen, 1977; Lahiri, 2002b). We capture the basic fact with a feature [+presupposed] projected from *know*:

12) John knew where Bill went.

The sentence in 12) presupposes that Bill went somewhere, and it presupposes that the knowledge is complete as evaluated both from the matrix subject’s perspective and the speaker’s PoV perspective.

In the case of *wonder*, the feature must be something that captures the phenomena of an “indirect question”, which means that it will be exempt from local Transfer and interpreted at a higher Phase:

11) John wondered where Bill went.

The case of the verb *think* is remarkable in that the complement clause cannot contain an overt wh-complementizer:

13) *John thought where Bill went.
14) John thought that Bill went.

*Think* precludes the meanings of *either* an indirect question or a presupposed interpretation. Overt expression of the wh-word is excluded in this environment.

Notice that the effect is limited precisely to where the higher verb (e.g. *think*) projects a PoV + opacity Feature Bundle onto a lower CP. The effect disappears where no CP is present, for example, with a PP. There, the automatic, default factivity returns, just as it does for a verb like *know*:

15) John thought *about* where Bill went.

This presupposes that Bill went somewhere.

In sum, we have a series of verbs: *know, wonder, think*, that take different kinds of complements (Égré, 2005; Lahiri, 2002b). Only some (like *wonder, ask*) take an indirect question. In addition, the verbs take different kinds of PoV. The complement under *wonder*, as in the cases of *think, believe, hope, pretend*, seems to be the subject’s PoV, and therefore opaque (Sharvit, 2002). Consider a case where Alice bought something that John mistakenly believes is a cake, but it is instead a hat in a hatbox. The sentences with *wonder* and *ask* allow John’s PoV in the lower clause, but not *know*:

16) John wondered/asked where Alice bought the lovely cake.
17) *John knew where Alice bought the lovely cake.

To point forward to the special case of African American English, and to keep the full range of data in our account, consider how AAE expresses indirect questions (Green, 2002). We discuss the fact that AAE allows inversion to occur in the lower clause of an indirect question:

18) John asked could he do it?

---

1 *Know* is not considered a traditional factive verb. Kiparsky and Kiparsky (1970) classified it as semi-factive on the basis of e.g. how it behaves under negation with first person subjects: I *don’t know that Bill went home* does not presuppose the truth of its complement, because it would create a contradiction. The topic deserves a paper in its own right, because it interacts in a complex way with discourse and focus. One can change the scope of the negation by stressing it:

He *doesn’t know* that Bill went home.

and follow with *because Bill never actually left*, and thus not presuppose the truth of the complement. But in that case, *know* is no different from a classical factive like *regret*:

He *didn’t regret* that Bill went home, because Bill never actually left.

The concept “presuppose” is part of a semantic system, not a necessary reflection of reality. In effect then, *know that* does presuppose the truth of the complement in its semantics, but that truth can still be challenged in a more complex discourse.

Nevertheless, the case in question with a wh-complement under *don’t know*:

John didn’t know where Bill went.

does not presuppose that Bill went anywhere. Later we show that the negative version *don’t know* behaves differently in AAE precisely because the presupposition is lifted.
In contrast, Mainstream English requires non-inversion, and a complementizer:

19) John asked if he could do it.

But if we consider the verb *know*, inversion in AAE is blocked just in case there is a presupposition of truth in affirmation sentences:

20) *John knew could he do it.

but not in the case of negation where that presupposition is lifted:

21) John didn’t know could he do it.

This implies that the auxiliary inversion is linked to an indirect, [−presupposition] choice. It is blocked when the verb entails a presupposition, as in 20) (thanks to Lisa Green for this observation). This fact will play an important role below. Table 1 presents these phenomena across verbs in summary.⁴

From this we see that the projection of a “feature” from the matrix verb to the complement is no simple matter. It is not that the projection of the feature “subject PoV” precludes wh- (because of wonder, ask). Projection of this PoV Feature Bundle disallows “closure” of the proposition (think).

This brings us to the last case that must be considered, namely the partial movement case in 10). There is plentiful evidence that the English-speaking child makes mistakes on 4) (repeated here):

4) How did she ask what to bake?

as if they treat it as a case of 10). What features do children lack that allows this misinterpretation, and how and when is it resolved? One possibility that we have pursued is that it is resolved at the same point in development as the Subj-PoV is resolved, but that seems to be empirically false: the error persists beyond the error on the opaque complement, until 6 or 7 years (de Villiers and Pyers, 2002). So we must dig deeper into features again.

There is an interesting controversy about how partial movement sentences like (10) are interpreted in languages that allow them, like German:

10) Was hat Hans gesagt wer gekommen ist?
What did Hans say who came?
 ’Who did Hans say came?’

Some speakers claim that this structure is only allowed, or interpretable, when what Hans says is “true” (in the weak sense that we use this word) (Herburger, 1994). This “true” interpretation has recently been called “speaker factivity” by Guerzoni and Sharvit (2007). Other speakers are less sure, since the higher was operates as a scope-marker and imposes opacity on the lower wh-word, hence dropping the factivity requirement. This tension follows predictably if there is a default interpretation at the Phase level, as occurs with adjunct clauses. If there is an overt wh-phrase and no indirect question feature, then transfer occurs. The uncertainty of adult judgments is just what one would predict when there is both a default interpretation, triggered by the overt wh-word, and a conflict with a projected feature.

Having considered each of the three cases, we now propose a working feature system, a Feature Bundle whose aim is to capture the intuitive facts and the acquisition stage. This will help us to specify what children at first apparently lack.

1.1. Building a Feature Bundle

The goal is to build a Feature Bundle that can reflect the complex changes in language acquisition. The first pivotal fact is that the matrix verb projects features onto its argument structure, including the complement. A formal feature captures the possible complement types under a verb:

Table 1

<table>
<thead>
<tr>
<th></th>
<th>Wonder/ask</th>
<th>Know/forget</th>
<th>Think/hope</th>
<th>Didn’t know</th>
</tr>
</thead>
<tbody>
<tr>
<td>PoV in complement</td>
<td>Subject</td>
<td>Speaker/universal</td>
<td>Subject</td>
<td>Speaker/universal</td>
</tr>
<tr>
<td>Wh complement</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Indirect question</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Presupposed/closed complement</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

⁴ An anonymous reviewer reminded us that verbs like *forget* take other complements that are not presupposed, such as the infinitival *John forgot to wash his socks* (see Perez-Leroux and Schulz, 1999; Roeper and de Villiers, 1994). Therefore, the properties of “presupposed complement” cannot simply be attached to matrix verbs, but interact with the tense of the complement in ways that take us beyond the scope of this paper. This further demonstrates the challenge that the child faces in acquiring the complement system.
We take [+presupposition] to be the default form carrying the general\(^5\) or speaker PoV, resulting in factivity.

Next consider a wh-word that begins in a lower clause. We assume that all such elements carry the feature [+wh], and that this feature attracts the element to the Edge of the Phase. However, only some wh-words carry [+Q] designating a question force. If there is a [+Q] feature on the [+wh] word, this feature will not be satisfied by movement to the Edge of the first Phase. It will continue to the higher clause where it will engage in a wh-scope partial movement configuration in the left periphery of the CP where the illocutionary force of a question, represented as [+Q] in the Force Phrase subpart of the left periphery, can be checked off (Rizzi, 2006).

Thus we have the following configuration under know for the sentence:

\[
\text{22) John knew what Bill did} \quad \text{t} \quad \text{Bill did} \quad \text{t}
\]

\[
\text{CP1–Force} \\
\text{[Assertion]} \\
\text{... VP} \\
\text{know} \\
\text{CP2} \\
\text{[+wh, + presupposition] = +factive} \\
\text{VP} \\
\text{| what} \\
\text{[+wh].}
\]

That is, the wh word begins in the lower clause and moves to check off the [+wh] feature in the medial CP. The [+wh] feature is present in Spec-CP as a formal device to allow cyclic wh-movement. Transfer to both phonological and interpretative interfaces occurs, with the default factivity carried by the feature [+presupposition]. The lower proposition is closed, presupposed, and it is not an indirect question.

Suppose instead that the element has both the features [+wh] and [+Q]. Then it moves up to the Force Phrase of the top CP, and we get:

\[
\text{23) What did John know t Bill did t} \quad \text{t} \quad \text{Bill did} \quad \text{t}
\]

\[
\text{[+Q]}
\]

Next consider the case of the verb think. The upper verb projects its own arguments, and they dictate properties of a lower clause. With a verb like think, the option of [−presupposition] requires some triggering experience for the child, some evidence of Subject-PoV (see J.G. de Villiers, 2004). Assume that Subject-PoV is associated with [−presupposition] in the verb’s Feature Bundle, which could be a feature of the UG representation of verbs.

\[
\text{24) What did John think t Bill did t} \quad \text{t} \quad \text{Bill did} \quad \text{t}
\]

\[
\text{[+Q]}
\]

\[\text{5 There are many lively questions here concerning whether there is a general PoV (Seat of Knowledge?) that is separable from Speaker PoV, but the debate takes us too far afield (Speas, 2000).}\]
The feature [+wh] attracts the wh-word to the medial CP, but then it must move higher to satisfy [+Q] and that moves it on to the top CP, leaving only a trace. The verb think does not license an indirect question. If the wh-word in the lower clause lacked [+Q], it would move only to the medial position and local transfer would occur, just as with know, closing off the clause and producing 25):

25) *John thought what Bill said.

In adult grammar, think has the feature [−presupposition] on its CP, dictated by Subj-PoV. This blocks transfer of the wh-element at the first Edge to interpretation/phonology. In the absence of such a feature, transfer is predicted to occur, and the child would “interpret” the medial trace with the lower verb.

Finally, consider the opposite case, of the verb wonder. Wonder licenses an indirect question, that we will mark with the abstract feature [−Q], to indicate that it is incompatible with [+Q] in the same location.

26) John wondered what Bill said

Finally, we come to the case of partial movement.

28) Was hat er gesagt was Bill gekauft hat?

‘What has he said what Bill bought t?’

In this way think and wonder are opposites, captured by the feature [−Q] on wonder.
because of the feature [–presupposition]. But the features are not completely satisfied until the wh moves covertly to the top CP to check [+Q], and a scope marker is generated in the top CP. If the feature [–Q] were not there, then wh would transfer to interpretation and the clause will be closed. We assume that different speakers of e.g. German might disagree on this reading precisely because the status of an indirect question under a verb like say is ambiguous. Consider the following case in English. John sees a woman walk in and mistakes her for Alice. In fact it is Sarah. John says, Alice arrived! The question is,

29) Did John say who arrived?

Most speakers hesitate between two possible readings: an indirect question, in which case the answer is yes, and a closed, true reading of the lower clause to which the answer is no. Perhaps the same uncertainty about the feature [–Q] is present in the German partial movement case.

Now take a child hearing a sentence with a medial wh (e.g. de Villiers et al., 1990):

30) How did the mother ask what to bake?

We assume the presence of the wh-word in the top Spec-CP is taken as evidence of [+Q], or question force. The child’s grammar considers the overt medial wh-word as evidence of phonological transfer, therefore undertakes interpretive transfer of wh at the first Phase and arrives at a default factive interpretation. The child then answers the medial question. Thus, the overt wh causes default factivity. The wh-feature in the higher clause attracts covert movement to the Force phrase, which then requires an “answer”, and what is available is:

31) What did the mother bake?

When the child gets the feature [–Q] on such verbs, this feature blocks local transfer to interpretation and [–presupposition] predominates despite the overt wh.

In sum:

a) [+Q moves [+wh, +Q] on to the top Phase (overtly in long distance movement or covertly in partial movement)

b) –Q obligatorily transfers [+wh] to phonology (indirect question)

c) –Q disallows transfer of [+wh] to interpretation and hence default factivity because it is always associated with [–presupposition]

d) In the absence of [–Q] or [+Q], [+wh] is transferred to both phonology and interpretation, i.e. overt in the medial CP and the clause is closed (the case of John knew what Bill did).

This representation is more highly articulated than most accounts: one could say it is on the Edge! We have done this deliberately to provoke work that will refine it or make a better competing proposal.

1.2. Trigger to adulthood

How does the child overcome the partial movement and incorrect long distance interpretation? When children acquire a PoV Feature Bundle on the embedded clause – and we have argued that the verb say plays a crucial role in this discovery (de Villiers, 2005) – then interpretation of opaque questions is delayed until the matrix clause as in adult English. One aspect of this feature projection of Subj-PoV to a complement is that children then have the representation they need to engage in explicit reasoning about false beliefs. Indeed, several convergent pieces of evidence suggest that acquiring the syntax of complement clauses is prerequisite for passing false belief explicit reasoning tasks in children, even when the false belief tasks are nonverbal (de Villiers, 2007).

The second point of interest is what children do with medial wh-questions; that is, cases in which there is an overt wh-complementizer in the sentence. In a sentence such as:

32) How did Bill say who played soccer?

The medial who blocks the long distance interpretation of the how, so an appropriate answer cannot be e.g. in his socks, but must be linked to the matrix verb e.g. he shouted, he whispered etc. Adults and children as young as four years of age disallow long distance interpretation of the how, but they differ in their responses (de Villiers et al., 1990). Young children are prone to answer the who. We see that the default interpretation at the first Phase boundary forces the child to treat the how as a scope-marker in a partial movement construction when asked such questions.

1.3. The AAE acquisition path

What light does AAE shed on this story? We predict that children who speak AAE versus MAE will have a very similar time course for development of the grammar of complex wh-questions and complementation, all else being equal (e.g. given similar socioeconomic environments).
In AAE, indirect questions allow inversion of the lower subject and auxiliary, unlike MAE, as mentioned above:

33) The girl asked can she go outside. (AAE)
34) The girl asked if she can go outside. (MAE)

This inversion is restricted to genuine indirect questions, so it is not possible in AAE with predicates of know:

35) *John knew can he go. (AAE)
36) John knew whether he can go. (AAE and MAE)

Therefore the presence of inversion in AAE subordinate clauses provides the child with an advantage in the comprehension of indirect questions, revealing that certain verbs take indirect questions. It cannot be mistaken for a direct assertion. In MAE there is no construction which appears uniquely as an indirect question.

Learners of AAE therefore receive clear syntactically unambiguous evidence of an indirect question form, whereas learners of MAE receive a different construction for matrix and subordinate yes/no questions. If this is correct, then AAE speakers have an advantage in recognition that certain verbs take indirect questions. In our Feature Bundle representation, once the child marks the feature \([-Q]\) which is forced by pure indirect questions:

37) John asked can he go.

then the other indirect option becomes visible and recognizable in the input:

38) John asked what he wanted.

This blocks the partial movement option for:

39) How did John ask what he wanted?

The consequence is that AAE learners will show “earlier” resolution of the medial answers than MAE learners because they have another option which allows Transfer of interpretation to the next phase for overt wh-expressions. In particular, those AAE speakers who use inversion of auxiliary and subject to mark indirect questions in an elicitation task, will be more advanced in their resistance of medial wh answers in a question comprehension task.

1.4. The case of SLI

The prediction for children with language impairment would be that both of the errors that occur in the course of normal acquisition, namely default factivity and answering the medial question, will be prevalent and persist longer than in the normal course of acquisition. We present data below from a large study of specific language-impaired children that fulfill these predictions.

1.5. Hypotheses

The data allow us to address the following hypotheses:

1. The first prediction is that AAE-speaking children and MAE-speaking children will both show a tendency to answer the long distance question as if the trace is interpreted with respect to only the lower clause. This will be a prevalent tendency in the young subjects of both dialects and will resolve at the same age for each dialect.
2. The second prediction is that both groups of AAE-speaking and MAE-speaking children will make the mistake of answering the medial (indirect) question, but that AAE-speaking children will abandon that error at a younger age than MAE-speaking children.
3. The third prediction is that children with specific language impairment will show prolonged errors of both these types on the wh-question tasks.
4. The fourth prediction is the AAE-speaking children will produce indirect speech reports in which the embedded clause has inversion of auxiliary and subject, whereas MAE speakers will rarely produce that form.
5. As a corollary of 2 and 4, the fifth prediction is that those AAE-speaking children who use auxiliary inversion in the embedded clause of their indirect speech reports will be less likely to answer the medial question than those who do not.

2. Data

The data for the present paper came from field-testing by The Psychological Corporation of the Dialect Sensitive Language Test (DSLT) (Seymour et al., 2000) the precursor to the Diagnostic Evaluation of Language Variation – Norm Referenced (DELV-NR) (Seymour et al., 2005). The DELV-NR is a new comprehensive assessment of 4 through 9-year-old children’s
syntax, semantics, pragmatics and speech production skills designed to be unbiased against speakers of AAE and related dialects. The DSLT was the field-testing version of the DELV. It contained all of the same subtests, but with many more items on each subtest than were retained in the final standardized version of the DELV, as it took almost two hours to administer. All of the testing was conducted by certified speech and language pathologists (SLPs) hired by the publishing company.

2.1. Subjects

The participants in the present study were 590 children aged 4–9 years who had taken all of the language subtests from the DSLT. There were 298 girls and 292 boys. 352 were AAE speakers and 238 were MAE speakers. In all, 398 of the children were typically developing and 192 were language-impaired with no other major cognitive impairments (30% of the AAE-speakers and 36% of the MAE-speakers). The children all had normal vision and hearing and normal range intelligence according to information reported to the field-testing SLPs from school records. The participating SLPs categorized the children as MAE- or AAE-speakers based on their own experience of the child’s speech as well as the child’s family and community background. A certified SLP at The Psychological Corporation also confirmed the children’s dialect status based on the presence or absence of distinctive morphological, phonological, syntactic and semantic features in the children’s verbatim responses to DSLT items recorded by the testing clinicians. The clinicians classified the participants “a priori” as typically developing (TD) or specifically language impaired (SLI) based on their standard practices of language assessment, defined by their school district. The SLI children were all receiving intervention services for language impairment.

All the children had also taken the DELV Screening Test (DELV-ST) (Seymour et al., 2003). This a 15 min assessment with two parts: a set of dialect-identifying phonological and morphosyntax items that contrast sharply between AAE and MAE, and a set of diagnostic items that are neutral with respect to dialect but have been shown to discriminate between TD and SLI children in prior research. The diagnostic items on the DELV-ST include dialect neutral morphosyntax, complex sentence comprehension, and repetition of non-word verbal stimuli that use only phonological elements that are common to AAE and MAE.

The 590 subjects were selected from more than 1000 participants in the DSLT field-testing. They were selected according to a combination of their performance on the DELV-ST and the SLPs’ categorization of them. Thus, to be classified as typically developing (TD) the children had to have been categorized by the testing SLPs as TD and in addition their performance on the DELV-ST diagnostic items had to fall into the “low risk” or “lowest risk” range of scores for their age. Correspondingly, to be classified as specific language impaired the children had to have been identified by the SLPs as SLI and their performance on the diagnostic section of the DELV-ST had to fall into the “highest risk” range of scores for their age. The 590 participants were therefore more strictly defined as TD or SLI than the overall field-tested sample of children.

2.2. Procedure

The Communicative Role Taking subtest of the DSLT elicited two examples of each of the following speech acts from the children, based on a two-picture sequence: reporting an event (what someone was telling someone else); making a request (what someone was asking someone else); and prohibiting an action (what someone was saying to someone else in a clearly negative situation). The present paper focuses on the responses produced for the request eliciting items. The results from the full subtest are described in P.A. de Villiers (2004) and in Johnson et al. (2005).

Fig. 1 shows one of the stimuli designed to elicit a direct or indirect question. The children were first shown the single picture on the left and told: “Look at what is happening here.” Then they were shown both pictures and the testing clinician pointed to the picture on the right and said: “What is the girl asking her mother?”

The Wh-Question Comprehension subtest in the Syntax Domain section of the DSLT used short narrative scenarios and a sequence of three pictures to elicit answers to a variety of complex wh-questions. This question comprehension testing procedure is described in de Villiers and Roeper (1996). The present paper is concerned with the children’s answers to wh questions containing a medial wh-word such as

40) How did the woman learn what to bake?
41) How did the boy ask what to draw?

Results from the full set of wh-question types tested in the DSLT data are presented and discussed in de Villiers et al. (2008). For example, one story goes:

“These children are planning a surprise party for their favorite teacher. Everyone was going to bring some food for the party. But the boy didn’t know what to bring. He asked the woman at the grocery store, «What shall I bring my teacher?» The woman told him his teacher loved bologna, so that’s what the boy decided to bring her.””

“Who did the boy ask what to bring?”
The particular stimuli used in the present study contained one “false” complement:

42) What did the mother say she bought?
There were also four cases in which there was a medial wh-complementizer that could occasion a "medial question" answer. These varied in whether they contained argument or adjunct questions (for full information on how these behaved, see de Villiers et al., 2008). The examples were:

43) Who did the little sister ask how to see?
44) Who did the boy ask what to bring?
45) How did the mom learn what to bake?
46) How did the girl ask where to ride?

2.3. Coding of the children's responses

The syntactic form of the children's responses on the Communicative Role Taking task were coded into four different categories (Johnson et al., 2005):

1. Direct speech: e.g., "Can I have some cake?" or "Who is the cake for?"
2. Indirect speech: e.g., "She is asking her mother if she can have some cake." or "Who the cake is for."
3. Partially indirect/mixed: e.g., "Can she go eat cake?" (where the response seems to have features of both direct and indirect speech forms)
4. Ambiguous: e.g., "Eat cake." (where the response doesn't fit clearly into any of the above three categories)

Mixed responses such as those in category 3 are ungrammatical in MAE, but are characteristic of AAE (Green, 2002), as discussed earlier.

On the single long distance question (42), the children’s answers were coded as:

1. Long distance answers: e.g. “She said she bought paper towels”, or “Paper towels”.
2. Lower verb only: e.g. “She bought cake” or “Cake”.

The second answer represents an error, as it is a failure of scope; the child may be adhering to a locality constraint in answering the trace in the lower clause only.
Children's answers to the complex wh-questions with medial wh-words on the Wh-question Comprehension subtest were coded as:

1. Short distance answers: e.g., “The woman in the store.” (answering the top verb)
2. Long distance answers: e.g., “His teacher.” (answering the lower verb)
3. Medial answers: e.g., “Bologna.” (answering the medial wh-complementizer)

Long distance and medial answers are incorrect in both MAE and AAE for those items in which the medial wh-word is a complementizer and not a real question in English.

3. Results

There were no significant differences between the AAE- and MAE-speaking children in their success on the Communicative Role Taking or the Wh-Question Comprehension subtests on the DSLT (or finally on the DELV-NR) (de Villiers et al., 2008; Johnson et al., 2005; Seymour et al., 2005). Nevertheless, these subtests showed good properties for an effective overall assessment of language development, delay or disorder. On both subtests there were highly significant effects of age and clinical status, so the items were sensitive markers of language development and also discriminated clearly between TD and SLI children (de Villiers et al., 2008; P.A. de Villiers, 2004).

Table 2 summarizes the findings for the three syntactic phenomena that are the particular concern of the present paper: the tendency of children to incorrectly answer the wh-trace with only the lower verb in sentences such as 42); the tendency to answer the medial wh-question (partial movement) in sentences such as 43) to 46); and the production of indirect question reports in which the embedded clause has inversion of AUX and Subj, a distinctive feature of AAE syntax (Green, 2002). The table compares the performance of TD and SLI children as well as the similarities and differences between AAE and MAE speakers across the different age groups.

### Table 2

Comparison of TD and SLI children and MAE versus AAE speakers by age group on the three particular Wh-question syntactic phenomena addressed in the present paper. The table shows the mean percentage of items on which the form in question was produced.

<table>
<thead>
<tr>
<th>Age group</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7–9</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lower verb answers to 42) “What did she say she bought?”</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAE</td>
<td>34.3</td>
<td>22.6</td>
<td>17.9</td>
<td>9.3</td>
<td>20.5</td>
</tr>
<tr>
<td>MAE</td>
<td>32.0</td>
<td>31.8</td>
<td>15.4</td>
<td>12.2</td>
<td>22.7</td>
</tr>
<tr>
<td>TD</td>
<td>32.2</td>
<td>14.4</td>
<td>11.4</td>
<td>3.7</td>
<td>15.1</td>
</tr>
<tr>
<td>SLI</td>
<td>36.4</td>
<td>49.1</td>
<td>30.4</td>
<td>22.0</td>
<td>34.4</td>
</tr>
<tr>
<td><strong>Medial Wh answers – partial movement errors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAE</td>
<td>38.6</td>
<td>30.9</td>
<td>21.4</td>
<td>15.5</td>
<td>25.8</td>
</tr>
<tr>
<td>MAE</td>
<td>46.7</td>
<td>37.9</td>
<td>26.7</td>
<td>18.7</td>
<td>32.1</td>
</tr>
<tr>
<td>TD</td>
<td>37.5</td>
<td>19.9</td>
<td>12.1</td>
<td>6.1</td>
<td>18.3</td>
</tr>
<tr>
<td>SLI</td>
<td>53.5</td>
<td>59.7</td>
<td>49.4</td>
<td>34.7</td>
<td>49.1</td>
</tr>
<tr>
<td><strong>AAE indirect question forms with inversion of AUX and Subj</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAE</td>
<td>25.0</td>
<td>32.1</td>
<td>39.0</td>
<td>40.6</td>
<td>35.0</td>
</tr>
<tr>
<td>MAE</td>
<td>3.0</td>
<td>6.0</td>
<td>7.7</td>
<td>5.8</td>
<td>5.6</td>
</tr>
</tbody>
</table>

The prediction was made that children diagnosed as SLI would show more and later errors of two types: the first is answering the wh-trace with the lower verb for 42) What did the mother say she bought?

Young children (4 and 5 years old) of both dialects had difficulty providing the appropriate answer when the lower clause involved a false statement, with about one third of the children answering the second clause only, answering what the mother bought, not what she said she bought. A quarter of the children with SLI continued to make this error at age six and seven (see Table 2). An ANOVA with the between subjects variables of age, clinical status and dialect revealed a highly significant effect of Age $F(5,566) = 5.0, p < .001, \text{eta}^2 = .042$ and a significant effect of Clinical Status $F(1, 566) = 22.9, p < .001, \text{eta}^2 = .039$. However, the child’s dialect had no effect $F(1,566) = .88, p = .35, \text{n.s.}$, nor did age $\times$ dialect. Thus the first prediction was confirmed.

The second predicted error was answering the medial question in a question such as:

45) How did the woman learn what to bake?
The medial answers were tallied across the items of this type and the results were subjected to an ANOVA. There was a strong effect of Age \( (F(5,565) = 8.53, p < .001, \eta^2 = .070) \) as well as Clinical Status \( (F(1,565) = 98.6, p < .001, \eta^2 = .149) \). Children with SLI persisted in answering the medial on a third of the items after age seven, when it was rare in typically developing children (see Table 2). Together with the results on the first type, these results confirm prediction 2, that SLI children will be delayed in both developments.

### 3.2. AAE patterns

However, some interesting differences in the detailed pattern of responding to the particular items on these two subtests did appear between AAE and MAE. One difference between the two dialect groups is confirmed on the pattern of errors on the medial wh-question forms in the Wh-Question Comprehension subtest. Specifically, the MAE speakers were more likely to incorrectly answer the medial wh (see Table 2). MAE speaking children made significantly more medial answer errors and persisted in making the error to an older age than the AAE speaking children \( (F(1,475) = 4.65, p < .03) \). This confirms Hypothesis 2.

On the Communicative Role Taking task responses to the items that elicit direct and indirect speech reporting an event or negative imperatives prohibiting an action should not vary by dialect between MAE and AAE, and indeed there were no differences between the two dialect groups on these items in the frequency of various direct and indirect speech forms (Johnson et al., 2005). However, for the items eliciting requests (e.g., *What is the girl asking her mother?*), adult AAE allows for a distinctive indirect speech form that is ungrammatical in MAE (Green, 2002):

47) “The girl asked can she have some cake?”

The AAE speaking children were significantly more likely than MAE speakers to produce these forms \( (F(1,574) = 50.93, p < .0001) \) (see Table 2). In fact these “partial indirect” or “mixed” forms predominated in the productions of the AAE children and almost never occurred in the responses of the MAE children. Altogether the AAE children produced 246 responses coded as “partially indirect/mixed” compared to only 40 tokens of this type by the MAE children. The MAE children produced indirect speech forms with a complementizer and no inversion:

48) “The girl asked if she can have some cake.”

This confirms Hypothesis 4.

The fifth hypothesis of this paper is that these two observations are related through the syntactic differences of the two dialects and especially the distinctive inverted indirect speech question form in AAE. What evidence might support such a relationship between the two phenomena? To test this we asked whether individual AAE speaking children showed a relationship between their responses on the two different tasks. A Chi² analysis was carried out on the 352 AAE speaking children in the study to examine the relation between the provision of any medial answers on the Wh-Question Comprehension subtest versus the production of any AAE “inverted” embedded question forms in the Communicative Role Taking task. The statistical relationship was significant \( (\text{Chi}^2 = 4.08, p < .04) \). When the children were classified as consistent (2 tokens or more) providers of medial answers versus consistent (2 tokens) producers of AAE embedded question forms the statistical relationship was more significant \( (\text{Chi}^2 = 7.99, p < .005) \). In both of these cases the relationship was an inverse one: Children who produced the AAE embedded question form were significantly less likely to answer the medial question in the Wh-Question Comprehension task.

The various hypotheses and outcomes are summarized in Table 3.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Observed result</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a. Younger child answers the wh-Q with the lower clause in 42): “What did the mother say she bought?”</td>
<td>Strong Age effect for this error.</td>
<td>Age: ( F(5,566) = 5.0, p &lt; .001 )</td>
</tr>
<tr>
<td>1b. No AAE vs. MAE dialect differences in the likelihood of this error.</td>
<td>No Dialect differences for this error.</td>
<td>Dialect: ( F(1,566) = .88, p = .35 )</td>
</tr>
<tr>
<td>2. AAE-speakers stop answering the medial wh earlier in development than MAE-speaking children.</td>
<td>Dialect effect in favor of the AAE-speakers, i.e. fewer medial wh errors.</td>
<td>Dialect: ( F(1,475) = 4.65, p &lt; .03 )</td>
</tr>
<tr>
<td>3. SLI children make more of both of these errors for longer in development.</td>
<td>Strong Clinical Status effects found for both lower clause error and medial wh error – SLI &gt; TD children.</td>
<td>Clinical Status effect: Lower clause error: ( F(1,566) = 22.9, p &lt; .001 ) Medial wh error: ( F(1,565) = 98.6, p &lt; .001 )</td>
</tr>
<tr>
<td>4. AAE-speakers produce many more indirect question reports with inversion of AUX and Subj as in adult AAE.</td>
<td>Most AAE-speaking children produced these forms; almost no MAE-speakers did.</td>
<td>Dialect effect: ( F(1,574) = 50.93, p &lt; .0001 )</td>
</tr>
<tr>
<td>5. AAE-speakers who produce inversion of AUX and Subj in indirect question reports will not answer the medial wh.</td>
<td>AAE children who consistently produced AAE embedded Q forms were significantly less likely to answer the medial wh.</td>
<td>Chi² = 7.99, ( p &lt; .005 )</td>
</tr>
</tbody>
</table>
4. Discussion

We began with a proposal about how children’s grammars might adopt an idealization about wh-movement in which the interpretation is completed one Phase a time, or locally. We used this to explain the common phenomenon in which young children learning English answer the medial complementizer as if it were a real question, and extended it to account also for the case where a wh-trace is treated as if it belonged only to the lower clause. This was accomplished by linking elements in a Feature Bundle projected from the higher verb to the lower complement.

The prediction followed that the first step in the abandonment of the locality restriction would arise with the trace, once evidence accumulated that there was a Point of View Feature Bundle on the lower clause. The presence of this subject PoV feature then forces deferral of interpretation until the top clause. We present a technical feature machinery to accomplish this and to separate out different wh-forms, direct and indirect, factive and non-factive. Children with SLI had difficulty with this machinery and persisted longer in local transfer. However, as expected by the theory, there was no difference between AAE and MAE speakers in the resolution of this aspect of grammar. Acquiring the PoV feature should be no different across these dialects, provided dialect is the only difference between the groups.

Given the array of possibilities for complements under different verb types, the natural conclusion is that sorting out what is an indirect question might take some time in language development. We made the prediction that AAE speakers would show precocity in this development because of a related feature of their grammars, namely, the presence of overt marking of indirect questions via inversion in the embedded clause. Acquiring the feature that differentiates verbs that take indirect questions from those that do not allows the child to resist the local interpretation of the medial wh-word.

The present data allowed a preliminary test of this hypothesis but there are many other aspects to explore, especially in regards to how children handle indirect questions in general. We are beginning to explore how children respond, for example, to the case of:

49) Did John say who arrived?

Would this also show a precocity in speakers of AAE? We expect children to initially answer the medial wh, and this error should disappear faster for speakers of AAE. But then what? Will children continue to regard the lower clause as having default factivity? Or will it be opaque, and allow John’s PoV? These issues we will explore next.

There are many unanswered questions in this arena about other languages as well. What happens in the course of acquisition of a language where partial movement is allowed? Does acquiring such a language prolong the child’s assumption that the lower overt wh means that the clause has been transferred to interpretation? We have begun to explore this with a colleague testing children who are native learners of Romani (Kyuchukov et al., 2007). We have used the opportunity here to test a small piece of an evolving story about the development of wh-movement by exploiting the dialect differences across these two English dialects. It is our hope that the results might also contribute to the larger domain, of opinion about children who speak African American English. Too often, the language of children speaking this dialect has been considered to constitute a deficit. Here we demonstrate an interesting advantage to one aspect of the grammar.

5. Conclusions

We attempt a feature checking account of wh-movement that posits a preference for local transfer in the child, with several consequences for interpretation of wh-questions. Local transfer to interpretation is first resisted in the case of a medial wh-trace, as a function of features dictated by the Subject-PoV of the matrix verb. Local transfer to interpretation is found to persist with overt medial wh until children acquired the feature marking indirect questions. We predicted and found that children with specific language impairment have considerable delay in resisting medial interpretations of indirect questions.

We also predicted and found that young AAE speakers were less likely than MAE speakers to wrongly answer the medial question. Furthermore, the extent to which AAE learners resisted the wrong answer was predicted by their responses on a task that elicited indirect questions. This finding reveals that answering the medial seems to be linked to establishing which verbs take indirect questions, a decision that is clearer in AAE.

References


