How Speakers Refer: The Role of Accessibility

Jennifer E. Arnold*
Department of Psychology, University of North Carolina

Abstract
One of the core components of language is referring, which requires the speaker to choose between expressions that are highly explicit (e.g., the UNC professor, or Peter), and reduced lexical forms (e.g., he). This paper reviews claims that this process is largely driven by the accessibility or salience of the referent, and the psychological processes that underlie these effects. Two classes of constraint are examined: (1) Discourse status, which has traditionally been identified as the determinant of referential choices and (2) Non-linguistic processing constraints that increase the use of explicit forms. These effects together support a modified version of the traditional claim that speakers choose referential explicitness so that the listener can identify the referent, and underscore the need for accessibility to be mediated by a non-linguistic representation.

1. Introduction

One of the most important parts of speaking is the ability to refer to things. You couldn’t say The house is on fire without being able to refer to the house. But successful reference involves multiple decisions. Should you say the house, the house with a red door, or simply it? People make these decisions rapidly enough to keep up with the pace of normal conversation, perhaps 50 times a minute. Moreover, they have to generate appropriate expressions. Consider the following examples:

1. Elsi called Sarah. Elsi asked lots of questions. Elsi told Sarah lots of funny stories too.
2. Jeannette and Mira saw each other at the conference in Germany. She was only there for three days.

In (1), the repeated Elsi seems redundant and inappropriate, whereas the pronoun she in (2) is ambiguous to the point of confusion. The question then is what drives judgments about appropriateness, and what processes accomplish this.

This paper examines the psychological processes underlying the decision about how explicit to make an expression. For any referent, potential expressions fall along a hierarchy of explicitness (Givón 1983; Ariel 1990; Gundel et al. 1993; see Almor and Nair 2007, for a review), ranging from semantically rich expressions (the professor who edited Topicality) to shorter terms (the professor; Givón), to pronouns (he) or even zeros, where the expression is dropped completely (Ø went home). I focus on English, but similar choices arise in other languages, some of which use zeros in contexts where English uses pronouns (e.g., Spanish pro-drop: Ø escribo = I write). This paper focuses on the process of reference production, i.e., how a speaker/writer chooses among alternate forms. However, it is likely that both production and comprehension processes are influenced by similar preferences, so I will occasionally refer to research on reference comprehension as well.

The traditional explanation for how speakers choose among alternatives is that they aim to provide enough information to allow their addressee to identify a referent (e.g.,
Chafe 1976, 1994; Givón 1983; Ariel 1990; Gundel et al. 1993). We can call this the “Communicative Goal of Referring” hypothesis. This review begins by examining how this view is supported by analyses of how referential form is constrained by the preceding linguistic discourse. I focus especially on the claim that some information is more accessible than other information, and how accessibility can be defined. I then examine recent evidence that referential decisions are also influenced by processing load, and consider what this means for the mechanisms of using accessibility. The discussion is limited to referential choices that maintain the same perspective. Outside the scope of this paper are decisions that reflect choices of perspective (e.g., Betty vs. the professor vs. the woman; Clark and Wilkes-Gibbs 1986; Isaacs and Clark 1987; Schober and Clark 1989), word retrieval (e.g., Dell 1986; Levelt 1999), or indirect reference (e.g., Nunberg 1977, 2004; Clark 1992; Ward 2004).

2. The Communicative Goal of Referring

It is plausible to assume that speakers want their references to be understood. To this end, they must provide enough information for the listener to pick out the referent (for specific formulations of this hypothesis, see Ariel 1990; Gundel et al. 1993; Chafe 1994). Such a task is not simple, because all references are ambiguous. Even a very specific reference like Jennifer Arnold can refer to more than one person (Google rphonebook alone lists over 100).

For one, the speaker must distinguish the intended referent from potential competitors (Olson 1970; Givón 1983; Ariel 1990). For example, in (2) the pronoun is insufficient. Experimental research has shown that speakers tend to provide more detailed descriptions when doing so avoids ambiguity, for example saying the white square in the presence of multiple squares (Olson 1970). Likewise, speakers of English tend to use pronouns more if there is only one gender-matched character in the context (e.g., Mike invited Zenzi to visit. She…), in contrast with situations where the pronoun would match more than one character, in which speakers are more likely to use a more specific expression (e.g., Mike invited Amit to visit. Amit….; Francik 1985; Arnold and Griffin 2007).

However, speakers do not always need to disambiguate. Consider the story in (3).

3. Tony and Simon went on vacation. Tony picked the location. He had been there many times before.

In the third sentence, it is obvious that he refers to Tony. Indeed, repeating Tony would be unnatural. Why is it not necessary to disambiguate between the two males here?

The traditional explanation appeals to the idea that some information in a discourse is more accessible, or salient than other information. When people communicate verbally, they keep non-linguistic mental representations of the characters and events in the discourse (Bransford et al. 1972; van Dijk and Kintsch 1983; Johnson-Laird 1983). This allows discourse participants to keep track of what has been mentioned, and helps the listener to link the new reference with the representation of that entity in their model (Clark and Haviland 1977). But these representations are not flat: some information seems easier to access, for example Tony in (3). This property has been described as accessibility, salience, focus, or prominence. Here I use the concept accessibility to refer to the property of information that makes it easier to access, independent of ambiguity considerations. Note that accessibility is a property of the information, and not words themselves. Consider a story like Ellen flew from Philadelphia to California. She stopped over in Mexico.
City on the way. We might say that She refers to Ellen. But the Ellen referred to is the person, not the word. We represent the entity Ellen non-linguistically, and this representation may be more or less accessible.2

Accessible referents tend to be pronominalized, whether or not the pronoun is ambiguous. By contrast, more explicit expressions, and possibly modification, are needed for less privileged references (e.g., Givón 1983; Gundel et al. 1993; Chafe 1994; Brennan 1995; Grosz et al. 1995; Arnold 1998). The concept of accessibility is a powerful one, but it raises questions about what exactly accessibility means, and how things become accessible. In any situation, some things seem intuitively more important, central, and easier to think or talk about, for example Elsi in (1). But the field has not converged on a single definition of accessibility, or the cognitive processes behind it.

Most commonly accessibility has been linked with a referent’s discourse status, which refers to the role an entity has played in the preceding discourse.

2.1. LINGUISTIC DETERMINANTS OF ACCESSIBILITY: THE EFFECTS OF DISCOURSE STATUS

Here I review four discourse properties that confer accessibility: Givenness, Recency, Syntactic prominence, and Thematic Prominence. One potential problem with using discourse status to explain accessibility is circularity: accessibility is used to explain why a pronoun is possible, but at the same time we explain what accessibility is by examining the situations where pronouns are possible. Nevertheless, the correlation of pronoun use with discourse features can be useful for two reasons. First, the discourse properties that cluster together in their constraints on reference form also tend to cluster for other linguistic decisions like word order (Bock and Irwin 1980; Arnold et al. 2000a; Birner and Ward 1998, in press) or accenting (Hirschberg 1993; Terken and Hirschberg 1994; Venditti and Hirschberg 2003). Second, after identifying what these features are, we can then ask what they have in common, which will begin to identify the cognitive processes involved.

2.1.1. Givenness
Speakers generally reserve pronouns for referents that have already been evoked in the discourse. By contrast, “new” things must be introduced with descriptions or names, as in (4) (e.g., Chafe 1976; Prince 1981).

4. Gregory refurbished a beautiful old Victorian.
   a. Given reference: He did a beautiful job.
   b. New reference: The neighbors were pleased.3

The term given usually is used for things that have been mentioned verbally, although information can also be evoked non-linguistically (e.g., visually; Clark and Marshall 1981). While the given/new contrast is useful as a heuristic, it is neither necessary nor sufficient for describing the conditions for choosing pronouns. On one hand, new information can sometimes be pronominalized, for example saying Did he win yet? as a conversation starter on election night (Greene et al. 1994). On the other hand, some given references are more likely to be pronominalized than others.

2.1.2. Recency
Reference to information mentioned in recent clauses is more likely to be pronominal than less recent information (e.g., Givón 1983). For example, the open diamonds in
Figure 1 illustrate how the rate of pronouns/zeros declines for reference to less-recently mentioned entities. Recency only operates on a clause-by-clause basis. Within a clause, the relative accessibility of entities is more strongly determined by syntactic and thematic prominence.

2.1.3. Syntactic Prominence
Some arguments of a clause tend to be perceived as more accessible than others, most commonly the grammatical subject (e.g., Brennan et al. 1987; Brennan 1995; Arnold et al. 2000b). For example, it is more natural to use a pronoun for the subject argument Tom, and repeat the non-subject Herb, as in (5a) than vice versa, as in (5b). This bias is based on the actual grammatical function, and not the semantic role (i.e., whether Tom was the inviter or the invitee). In English the subject is often the first-mentioned referent, but research in other languages has suggested that both order of mention and grammatical subjectionhood independently influence referential preferences (Järvikivi et al. 2005; Kaiser and Trueswell in press). Other syntactic structures that are considered prominent are the focus of clefts (Arnold 1998, 1999; Almor 1999; Cowles et al. 2007; Foraker and McElree 2007), or the syntactic topic in languages like Japanese (e.g., Walker et al. 1994).

5. Tom invited Herb to go on a bike ride. / Tom was invited by Herb to go on a bike ride.
   a. He asked Herb to bring the snacks.
   b. Tom asked him to bring the snacks.

Syntactic prominence can be considered a property of the referent. That is, once Tom is mentioned in subject position, the non-linguistic representation of Tom reflects the
accessibility of this position. But this effect is modulated by the syntactic position of the subsequent referring expression as well. Reduced forms are more likely if the referring expression is in the parallel syntactic position of the following clause. That is, the preference for a pronoun in (5a), \textit{He asked Herb}, should be stronger than if it had been \textit{Herb asked him} (Arnold 1998, 2003; Arnold et al. 2009; see Chambers and Smyth 1998 for relevant evidence from comprehension). Likewise, while a parallel object pronoun is acceptable to refer to \textit{Herb} in (5b), it would be infelicitous as a non-parallel subject pronoun (\textit{He asked Tom}...).

2.1.4. Thematic Prominence and Coherence Relations
Discourse accessibility has also been claimed to be sensitive to an entity’s semantic role. For example, in transitive events with Stimulus and Experiencer roles, speakers and writers prefer to use pronouns for reference the Stimulus, once syntactic prominence is controlled (Stevenson et al. 1994; see also Garvey and Caramazza 1974). This parallels comprehenders’ biases (e.g., Stewart et al. 2000), for example to interpret \textit{she} as referring to Laura in (6) and Hannah in (7).

6. Experiencer-Stimulus: Hannah admired Laura enormously because she…
7. Stimulus-Experiencer: Hannah impressed Laura enormously because she…

These effects critically depend on assumptions about the coherence relation between the two utterances (Stevenson et al. 1994; Arnold 2001; Kehler 2002; Kertz et al. 2006; Rohde et al. 2007). In (6) and (7), the word \textit{because} forces the assumption that the second clause provides a cause for the event in the first. This leads to the inference that the pronoun refers to the referent who is more likely to be the cause of the preceding event (Brown and Fish 1983). By contrast, continuations like \textit{so} lead to different biases (Stevenson et al. 1994). Even in sentences without \textit{because}, there is still an intended coherence relation between the two utterances (Arnold 2001; Kehler 2002), and listeners’ assumptions about coherence influence their interpretation.

There is evidence that coherence relations also influence speakers’ referential choices, mostly from studies using the story-completion methodology. Participants are given a story fragment (e.g., \textit{John got the manuscript from Aparna},) and asked to add another line to the story (e.g., \textit{He read it}). For example, Arnold (2001) examined events of transfer, like \textit{Rosemary accepted an invitation from Craig}, or \textit{Wind sold the couch to Wally}. The goal role was either in subject position (e.g., \textit{Rosemary}), or object-of-preposition position (e.g., \textit{Wally}). When participants continued the story with the subject character, they almost always used pronouns. But when they continued with the non-subject character, they used pronouns more often when it was a goal than a source. The increased accessibility of goals was carried by those items in which they continued the story with information about the events that followed the stimulus event (see also Stevenson et al. 1994; Rohde et al. 2007). Thus, some elements become more accessible as a result of semantic coherence, in combination with syntactic prominence effects.

2.2. CALCULATING ACCESSIBILITY

In sum, recent and prominent mention in the discourse correlate with speakers’ referential choices. These are said to influence a referent’s accessibility. But so far all we have is a grab–bag of discourse features that seem to matter. The real question is why? Why does this particular set of properties confer accessibility? Three suggestions in the literature...
(not mutually exclusive) are that accessibility accrues to those entities that are (1) topical, (2) predictable, and (3) attended.

2.2.1. Accessibility as Topicality

Intuitively, discourses and sentences seem to be “about” some things more than others (Reinhart 1982). This notion has been formalized as the linguistic category topic, which has been used to explain phenomena like intonation (Halliday 1967) and word order (see Birner and Ward in press). Topicality has also been proposed to increase the likelihood of using underspecified referential forms, like pronouns (e.g., Broadbent 1973; Sanford and Garrod 1981; Ariel 1990; Gernsbacher 1990). The accessibility of sentential topics squares well with evidence that grammatical subjecthood confers accessibility, as subjects are considered a topical position (cf. Chafe 1976). Topics can also be identified at a discourse level, for a given discourse segment (e.g., Grosz and Sidner 1986; Asher 2004; Kehler 2004).

However, topic as an explanatory category is problematic. This is especially true if we assume that each sentence has to have a single topic (Reinhart 1982). In a simple sentence, like Andy brews beer, is the topic Andy? beer? brewing? One way around this problem is to treat topicality as a continuous property of all discourse entities (Givón 1983). Givón measures topicality with textual properties that reflect previous importance in the discourse (recency), upcoming importance (persistence), as well as ambiguity of potential expressions, which significantly broadens the concept of topicality.

A formal definition of topic appears as Centering Theory’s center of attention (Grosz et al. 1995; Walker et al. 1998a; see also Grosz and Sidner 1986). Centering is a theory of discourse coherence and reference form. The idea of topic appears in Centering as what they call the backward-looking center (Cb). This is the referential expression that refers to something from the previous clause, thereby connecting backward with the previous discourse. If there is more than one expression that refers to something previously mentioned, the Cb is the one that refers to the highest-ranked entity from the previous clause. All nominal entities in each clause (which are called the forward-looking centers, or Cfs) are ranked according to their grammatical functions: Subject > Object > Oblique (Brennan et al. 1987). Centering posits that the discourse will be incoherent if the Cb is not pronominalized and something else is.

<table>
<thead>
<tr>
<th>Cb</th>
<th>Cf list</th>
</tr>
</thead>
<tbody>
<tr>
<td>8a.</td>
<td>Susan saw Lyn at a conference.</td>
</tr>
<tr>
<td>b.</td>
<td>She hadn’t seen her for a long time.</td>
</tr>
<tr>
<td>c.</td>
<td>Lyn asked Susan if she had seen Barbara.</td>
</tr>
<tr>
<td></td>
<td>Ø [Susan, Lyn, conference]</td>
</tr>
<tr>
<td></td>
<td>Susan [Susan, Lyn]</td>
</tr>
<tr>
<td></td>
<td>Susan [Lyn, Susan, Barbara]</td>
</tr>
</tbody>
</table>

Thus, the center (i.e., the topic) is determined by the ranking of the grammatical functions of the previous clause. But in principle any other property of the text could also be relevant (see papers in Walker et al. 1998b).

2.2.2. Accessibility as Predictability: The Expectancy Hypothesis

A related proposal is that discourse features influence accessibility by providing information about the predictability of upcoming reference (Arnold 1998, 2001, 2008; Arnold et al. 2007a,b; Arnold and Tanenhaus in press). Accessible entities are those that are relatively likely to be mentioned in the current utterance – i.e., those with relatively high expectancy.

The role of expectancy is based on evidence that prominently mentioned entities are more likely to receive subsequent mention. That is, the things that correlate with
pronoun use also correlate with an increased likelihood that something will be mentioned at all, independent of reference form (Arnold 1998). Figure 1 illustrates how recency correlates with both how you refer (i.e., pronouns vs. names, plotted by open diamonds), and what you refer to (i.e., the frequency of reference overall, plotted by filled squares). For example, speakers are most likely to continue talking about things mentioned in the previous clause, and the likelihood of this referential event correlates with a high tendency to use pronouns for reference to entities that were mentioned in the previous clause.

Analyses of written and spoken corpora have also found that speakers mention the referent of the preceding grammatical subject more frequently than the referent of the preceding object. They are more likely to produce parallel reference (subject-to-subject or non-subject-to-non-subject), after controlling for the general subject-reference preference. They are more likely to refer to the focus of clefts than the non-focus. And they are more likely to refer to certain thematic roles, like goals (over sources). Those same patterns that increase a referent’s expectancy also increase the likelihood that the speaker will use a less-specified referent, like a pronoun. These correlations hold for both pronoun use in English and the use of zero forms in Spanish and Mapudungun (Arnold 1998, 2003).

Expectancy is consistent with claims that topicality influences accessibility. The likelihood of re-mention could be construed as one measure of topicality. But crucially, expectancy focuses on what is likely to be topical, not what was already topical (similar to Centering’s forward-looking centers, and to Givón’s (1983) concept of persistence). Expectancy also accounts for the fact that new information can sometimes be more accessible than given, as when the speaker is disfluent (Arnold, Fagnano, & Tanenhaus, 2003; Arnold et al. 2004a), or when describing an unfamiliar object (Arnold et al. 2007b).

Under the communicative goal of referring, a plausible mechanism for expectancy is as a mechanism for discourse participants to coordinate accessibility. Expectancy describes how easily the comprehender will be able to retrieve the referent. Speakers could thus calculate expectancy as an estimate of accessibility to the listener. But to the extent that expectancy is calculated from public and shared discourse information, it may be heavily influenced by speakers’ own knowledge of the discourse (Arnold 2008; cf. Bard and Aylett 2004).

2.2.3. Accessibility as Attention
A third conception of accessibility (which is also consistent with both topicality and predictability proposals) is that it is related to attention – i.e., the accessible entities are those that the discourse participants are devoting more attention to. This idea underlies Gundel et al.’s (1993) Givenness hierarchy, which links each of its six cognitive states to attention. The highest-ranked is the discourse focus, defined as the entity that is in focus of attention. Likewise, Centering models the relationship between attentional state and discourse coherence, where the basic unit is the backward and forward-looking center(s) of attention (see also Bower and Morrow 1990).

Note that the focus of attention in discourse is not the same as the concept of “focus” in information structure. Within an utterance, information is often divided between the topic or theme (usually old information), and the focus (usually new information; e.g., Chafe 1976; Vallduví 1993; see also Birner and Ward in press). Attentional focus is more like topichood than informational focus, although both can lead to the expectation that the information will be topical in the next utterance (Arnold 1998, 1999).
Nevertheless, the exact relationship between attention and accessibility has not been demonstrated. It appears that entities with a prominent discourse status draw the attention of the speaker and listener, which is likely to result in a stronger, more accessible representation of that entity in the discourse model (Foraker and McElree 2007). An open question is whether the reverse direction of causality obtains, i.e., whether attention itself increases the accessibility of discourse entities, independently of their role in the discourse. In the next section of this paper I report evidence that undivided attention helps maintain accessibility. But can attention to an entity actually increase its accessibility? What little evidence there is on this question comes from comprehension. Arnold and Lao (under review) used an exogenous visual capture cue to draw listeners’ visual attention to one region on the screen before the story started, and found that the attended character received increased consideration as the referent of an ambiguous pronoun. However, the effect was transitory. Thus, it seems likely that attentional mechanisms subserve the calculation of discourse accessibility, but accessibility is not isomorphic with attention.

2.3. SUMMARY: THE COMMUNICATIVE GOAL OF REFERRING HYPOTHESIS

The linguistic context of the discourse strongly constrains referential form choices. This is consistent with the idea that speakers design expressions for interpretability. Linguistic information is, in principle, available to all discourse participants. Therefore, things with prominent discourse statuses can be assumed to be accessible to the listener, making less-specified referents accessible.

It is less obvious why speakers ever choose the less-specific forms. Wouldn’t it be safer to always use proper names? There are several explanations in the literature. One is that less-specific forms are more efficient, as embodied in Grice’s maxim of quantity: Make your contribution as informative as required, but not more so (Grice 1975). A second view is that reference form types are specialized for referential situations. Ariel (1990) suggests that the referential form is a “marker” for the discourse status of the referent, which helps listeners identify the location of the referent in their mental representation. This idea is supported by evidence that reading is slowed when a repeated name is used for a highly accessible referent (Gordon et al. 1993; Hudson-D’Zmura and Tanenhaus 1998).

In either case, the communicative goal hypothesis works well as an ideal. But we have said little about the processes used to represent accessibility, or how they trigger production choices. The next section considers how non-communicative processing constraints influence reference form.

3. Processing Load Influences on Reference Form

In order to better understand the processes by which accessibility impacts reference form, we can examine how the above-mentioned effects of accessibility are changed under conditions of cognitive load. It would not be surprising if reference production, like other cognitive tasks, were impacted by processing load, for example distraction by a secondary task. A more interesting issue is the nature of load effects. One possibility is that speakers will become less consistent under load, perhaps because they have fewer resources to calculate the discourse conditions. A second possibility is that speakers will tend to choose more semantically general terms, like pronouns, perhaps because processing load makes it difficult to retrieve more specific lexical items. A third possibility is that
speakers will tend to choose more specific forms, like names and descriptions, perhaps because a decrease in cognitive resources decreases the accessibility of all representations.

Here I summarize four effects of processing load, all of which support the last possibility: processing load (even mild) increases the use of explicit forms.

3.1. DISFLUENCY

Speakers are often disfluent, e.g., saying “um,” “uh,” or repeating words. Disfluency is one of the most direct indicators that the speaker is experiencing processing load. This load is often related to discourse or utterance planning, although it could reflect other sources of difficulty (e.g., Goldman-Eisler 1968; Beattie 1979; Siegman 1979; Fox Tree and Clark 1997; Bortfield et al. 2001).

References in disfluent utterances are less likely to be pronominal than references in fluent utterances. For example, in Arnold et al.’s (2009) analysis of spoken narratives, adolescents (both those with and without autism) were more likely to use a pronoun in fluent utterances (81%; SE: 1%) than disfluent utterances (68%; SE: 2%; Arnold et al. 2009; see also Arnold and Griffin 2007). This analysis included all references to the three main characters (Sylvester, Tweety, and Granny) that occurred in subject position, regardless of the discourse status of the character being referred to. However, it is important to note that references to brand-new5 (previously unmentioned) entities were never pronominal in this dataset, regardless of whether the speaker was disfluent. This suggests that the effect of disfluency is not to simply make choices more “noisy,” or make speakers less reliant on the discourse context, but rather to decrease pronoun use overall.

3.2. PLANNING LOAD

Even in fluent utterances, speakers need to devote processing resources to planning and executing their utterances. Although some planning occurs incrementally, as the utterance unfolds (Griffin and Bock 2000; Griffin 2003), there is also evidence that some planning occurs utterance-initially (Ford 1982; Clark and Wasow 1998). As longer utterances are likely to incur more processing load than shorter utterances, we might predict that the grammatical subjects should be pronominal more often for short than long utterances. This is indeed what Arnold et al. (2009) found: 84% (SE 1%) pronouns for clauses with one to six words, and 76% (SE 2%) pronouns for clauses with 7+ words.

3.3. COMPETITION

Arnold and Griffin (2007) had participants tell a story based on a two-panel cartoon. The first panel pictured either one character, or two different-gender characters. Participants repeated the first line of the story (e.g., Daisy went for a boat ride {with Mickey} on the lake). The second picture showed one character doing something (e.g., Daisy rowing away), either with the second character or without. Participants generated another line for the story (e.g., Daisy left Mickey behind; or She rowed into the sunset). Pronouns were more common in the one-character than two-character stories. This could not be attributed to discourse status or ambiguity avoidance, as the referent was always the subject of the first sentence, the thematic content was the same (other than the presence of the second character), and the second character was always of a different gender, making any pronoun unambiguous. We concluded that the mental representation of the second character consumed resources, lowering the accessibility of the first character.
3.4. MEMORY/ATTENTION LOAD

Speakers are also less likely to produce pronouns when performing a secondary task. In Babwah (2008), participants gave pairs of instructions so a partner could move pictures on a computer, e.g., Click on the broom. Now move {it/the broom} to the cow. In the distraction condition, the computer beeped one to four times prior to the second instruction, and the speaker clicked on a button to indicate the number of beeps heard. Some participants never used pronouns at all. But those who did were more likely to do so on the no-distraction than distraction items (for a related finding, see Griffin and Arnold 2008).

As the above findings demonstrate, speakers are more likely to use pronouns when their processing resources are not taxed by utterance planning, discourse representation, or a secondary task. It is striking that these diverse manipulations all decreased pronoun use.

Similar evidence comes from research on the pronunciation of words, which tends to be shorter and less clear in many of the same discourse contexts that support attenuated lexical choices (e.g., Brown 1983; Fowler and Housum 1987; Terken and Hirschberg 1994; Jurafsky et al. 2001; Bard and Aylett 2004). Likewise, production load results in reduced pronunciations. First, the words surrounding disfluencies tend to be longer than fluent segments (Fox Tree and Clark 1997; Jurafsky et al. 1998; Shriberg 1999; Bell et al. 2003). Second, Christodoulou (2009) manipulated difficulty by asking participants to refer to either familiar objects, or complex unfamiliar designs. The carrier phrase, e.g., Click on the blue, was longer on average for unfamiliar referents (even after disfluent items were excluded). The color word (e.g., blue) also had a higher average pitch for unfamiliar items, indicating increased acoustic prominence.

However, contrasting findings are reported by Almor et al. (1999). They found an increased use of pronouns for individuals with Alzheimers, a condition that is characterized by diminished working memory capacity. The difference may be explained by Almor et al.’s suggestion that individuals with Alzheimers have an impaired ability to retrieve the intended name or description. Thus, working memory deficits decreased the availability of the lexical items themselves, resulting in pronouns even for less-accessible representations. By contrast, the cognitive load effects described above are assumed to influence non-linguistic representations, making them less accessible.

3.4.1. Mechanisms of Using Accessibility

The evidence reviewed so far demonstrates that there are two classes of constraint on reference form: (1) textual discourse status constraints and (2) processing load constraints. What kind of production system could support both?

We can consider (and reject) the possibility that processing load simply reduces the speaker’s ability to calculate discourse status. This view would allow discourse accessibility to be modeled purely in terms of discourse status (via expectancy, topicality, or something else), without reference to a mediating representation like activation. However, this view does not explain why processing load consistently increases the explicitness of referential forms, as opposed to simply making performance more noisy.

A more likely possibility treats accessibility as a gradient property of non-linguistic discourse representations, modeled as activation (Arnold and Griffin 2007). Discourse properties can either boost or dampen a representation, for example increasing activation if the referent occurs in subject position. The next time the speaker refers to this entity, a pronoun is acceptable only if the activation passes a particular threshold. If cognitive tasks compete for available resources (e.g., Baddeley and Hitch 1974), processing load will
decrease available resources for maintaining activation, and thus increase the use of explicit forms. An open question is whether all activations drop equally or not. This interpretation suggests that accessibility is not entirely a relative phenomenon (i.e., we don’t simply pick the “most accessible” entity to pronominalize), but rather is at least partly dependent on absolute measures of accessibility.

If we assume that accessibility is represented in terms of activation, we have to assume that this activation is represented separately from whatever activation might reflect the mere selection of that entity as a referent. That is, if a speaker decides to refer to Liane, the decision itself might involve activation of the non-linguistic representation of Liane. Yet this does not make a pronoun appropriate. Thus, accessibility-related activation must exclude knowledge of the intended referent.

4. Conclusion

Speakers choose pronouns (and other underspecified forms) when the referent has a prominent discourse status, but also when they are experiencing cognitive load. This pair of findings suggests that while the Communicative Role of Referring Hypothesis is largely correct, it operates on the basis of underlying discourse representation. The activation of these representations can be influenced by more than purely communicative considerations, including processing load.

The traditional view is that speakers choose referential forms on the basis of interpretability. This requires speakers to engage in audience design, considering what is currently accessible to someone else – at the very least a generic addressee (e.g., Schober and Brennan 2003). As linguistic information is public, it is at least plausible that speakers use this to estimate accessibility to their addressee. But processing load effects are inherently internal to the production system. What this suggests is that the production system may achieve the goal of communication, but the mechanism itself involves calculations that are not privy to representations about one’s interlocutor (for further discussion of audience design in reference production, see Arnold 2008). Similarly, speakers mostly achieve the communicative goal of distinguishing the referent from alternatives in the context, but their ability to do so fluently is dependent on their attention to potential competitors (Ferreira et al. 2005; Brown-Schmidt and Tanenhaus 2006; Brown-Schmidt and Konopka 2008), and they sometimes provide more information than is necessary from the listener’s perspective (Nadig and Sedivy 2002; Wardlow Lane et al. 2006).

In any case, the effects of accessibility on reference production are mediated by the non-linguistic representations of discourse entities. While discourse status strongly influences the accessibility of these representations, non-linguistic processing load can moderate their strength as well. Thus, speakers’ decisions about referential forms stem from both discourse and cognitive resource-related constraints on the accessibility of non-linguistic mental representations.

Short Biography

Jennifer E. Arnold is an Assistant Professor of Psychology at the University of North Carolina at Chapel Hill. She does research on the mental processes involved in language comprehension and production, in particular how people use and understand references. She has published numerous papers on both the processes by which speakers choose between forms like “she” or “the woman,” and how listeners identify the referent of these expressions in real time. Her research is currently funded by the National Science
Foundation. She received her Ph.D. in linguistics from Stanford University in 1992, and did postdoctoral training in cognitive science and psychology at the University of Pennsylvania and the University of Rochester.

Acknowledgement
Thank you to Alex Christodoulou, Jason Kahn, and the students in my fall 2009 psychology of language class for feedback on this manuscript. This work was partially supported by NSF grant BCS-0745627.

Notes
* Correspondence address: Jennifer E. Arnold, CB #3270 Davie Hall, Rm. 337B, Chapel Hill, NC 27599-3270. E-mail: jarnold@email.unc.edu

1 Conversational speech is spoken at about 200 words per minute (Krause and Braida 2002); if 25% of all words are nouns, then we produce roughly 50 nouns per minute.
2 Conceptual and linguistic representations are associated with each other (e.g., Dell 1986; Levelt 1989), so conceptual accessibility may also increase or decrease the accessibility of associated lexical items. Lexical accessibility influences related processes, like acoustic realization of references (e.g., Bell et al. 2009). But as frequent pronouns are always more lexically accessible than any other form, this does not account for the majority of referential form choices.
3 The referent “neighbors” could arguably be classified as inferable instead of new (Prince 1992), as houses typically have neighbors, but both would function similarly with respect to pronominalization.
4 A language spoken by the Mapuche people, an indigenous people native to Chile and Argentina.
5 For our current purposes, I am using “brand-new” to include only references to characters that had never been mentioned before at all. This should not be confused with the coding used in Arnold et al. (2009), in which “unmentioned” was defined within a narrative episode.
6 Some entities are stereotypically gendered (e.g., fireman) rather than categorically.

Works Cited


Appendix: In-Class Exercise

This is an in-class experiment that can be conducted as a demonstration, or as the basis for a written assignment. I have conducted this with undergraduate students in psycholinguistics, cognitive psychology, and cognitive science classes over the last 12 years. The results always show the same pattern. A version of this experiment was published as the norming experiment in Arnold et al. (2000b).

This activity demonstrates how both discourse accessibility and ambiguity avoidance influence the production of referential expressions. Students provide a natural continuation to story fragments like The {mother/widower} fired the nanny at the dinner table. For example, they might write He didn’t have any more money to pay her, or The nanny was caught beating the children. They then analyze the form of the subject expression in their continuations (pronoun vs. name).

Discourse accessibility is measured in terms of order-of-mention/subjecthood. Ambiguity is manipulated within each item, which have two characters of either the same or different gender. This allows us to test three hypotheses: 1) pronouns are used more often for reference to the first-mentioned character (the widow/widower); 2) pronouns are used more often in the different-gender condition, and 3) the two effects interact, such that the gender effect is stronger for second-mentioned references.

MATERIALS AND DESIGN

Twelve single-sentence stories comprise the experimental items. Each has two characters, which are always gender-specific descriptions, like waitress, king, nun, princess.6 These items are divided into two lists, such that each has six same-gender and six different-gender items. The experimental items are mixed with eight fillers.

PROCEDURE

Students are given the questionnaire ahead of time, without any explanation of what it is for. If conducted in class, allow about 20 minutes. Then the students turn into “experimenters,” and the hypotheses and coding procedures are explained. Students code the responses to the critical items (either their own or someone else’s), by identifying the grammatical subject in their response. If it refers to one of the two human characters, they identify the referent (1st vs. 2nd character), and whether the expression is a pronoun or description. Class results are tabulated in terms of the percentage of pronouns used for the four categories (1st character/same gender; 1st character/different gender; 2nd character/same gender; 2nd character/different gender).

EXPECTED RESULTS

Results are expected to support all three hypotheses: (1) a main effect of order-of-mention (note that in this experiment this is confounded with parallelism), (2) a main effect of gender, and (3) an interaction between the two. Numerical averages generally mimic the pattern reported in Arnold et al. (2000b): Same gender/first-mention: 95%; Same gender/second-mention: 37%; Different-gender/first-mention: 95%; Different-gender/second-mention: 76%.
SUGGESTED VARIANTS

1. Write your own stimuli to examine the effects of other discourse variables, e.g., contrasting different verb types, animate vs. inanimate referents, or adding another sentence that mentions one of the two characters.
2. Examine the effects of processing variables by asking participants to complete the questionnaire under different circumstances (e.g., while also repeating numbers), or by comparing individuals from different categories, for example gender (see Arnold, 2010) or college major.

STIMULI

These stimuli are based on stimuli from an unpublished experiment carried out in collaboration with Amit Almor.

Experimental

1. The nun/priest whispered to the monk in the courtroom.
2. The fireman saved the man/woman from the burning building.
3. The girl/boy approached the policeman in the shopping mall.
4. The landlord/landlady confronted the woman on the stairs.
5. The waitress/waiter tripped the lady by the restroom.
6. The tailor/seamstress measured the pope in the dressing room.
7. The prince/princess visited the sorceress in the damp cave.
8. The maid/butler served the President in the Oval Office.
9. The headmistress praised the schoolgirl/schoolboy in the ceremony.
10. The salesman/saleslady overcharged the duchess in the antique shop.
11. The mother/widower fired the nanny at the dinner table.
12. The bride/groom embraced the maid of honor after the wedding.

Fillers

1. The workers huddled in a group with the supervisor.
2. The chef remembered the recipe was easy to make.
3. The defendant examined by the lawyer turned out to be unreliable.
4. A group of kids lined up at the ice cream truck.
5. The audience waited anxiously for the play to begin.
6. The senior citizen’s center held a New Year’s Party this year.
7. The prisoner read the trashy newspaper with fascination.
8. The building inspector wearily looked into the attic of the house.