Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 57 (2007) 49–64
Journal of Memory and Language
www.elsevier.com/locate/jml

Children's overtensing errors: Phonological and lexical effects on syntax

Joseph Paul Stemberger *

Department of Linguistics, University of British Columbia, E270-1866 Main Mall, Vancouver, BC, Canada V6T 1Z1

Received 5 May 2005; revision received 16 February 2007
Available online 6 April 2007

Abstract

Overtensing (the use of an inflected form in place of a nonfinite form, e.g. **didn't br**o**ke* for target *didn't br**ea**k*) is common in early syntax. In a ChiLDES-based study of 36 children acquiring English, I examine the effects of phonological and lexical factors. For irregulars, errors are more common with verbs of low frequency and when phonological processing biases favour the past-tense form relative to the base form (vowel dominance and the consonant addition bias). For regulars, errors are more common when the inflected form ends in a rime that can occur in monomorphemic forms in English (which children have had independent practice with) than when the rime is found only in inflected forms. Results demonstrate that default patterns can be subject to lexical frequency effects. Results are compatible with a particular conceptualization of competition (the integrated multiple competitor approach), whether connectionist or symbolic.

Keywords: Morphology; English past-tense; First language acquisition; Overtensing; Phonology–morphology interactions; Frequency effects; Regularity; Language production; Dual route models; Connectionist models; Symbolic models

When accessing words during sentence production, a speaker is often required to encode morphosyntactic information (such as tense and number) in the form of morphology, which is expressed as a set of phonological elements. The past few decades have seen intense research on a number of aspects of morphology (e.g., Bybee & Slobin, 1982; Marchman, 1997; Marcus et al., 1992). The most heated debate has focused on the issue of the mechanisms associated with regularity: Do speakers use different mechanisms for irregular past-tense forms (such as past-tense *broke* and *sang*, cf. base *break* and *sing*), which are relatively few in number and follow many idiosyncratic patterns, than for regular forms (such as *walked* and *needed*, cf. base *walk* and *need*), which are many in number and follow a single pattern? Proponents of two-mechanism approaches argue that data require different mechanisms for irregulars than for regulars, and that there in principle can be no lexical frequency effects on default patterns that are not stored in the lexicon (e.g., Pinker & Prince, 1988; Pinker & Ullman, 2002). Proponents of single-mechanism approaches argue that the data do not compellingly demonstrate the need for different mechanisms, and that lexical frequency should in principle arise even with default patterns (e.g., Rumelhart & McClelland, 1986; Stemberger, 2004). One difficulty in evaluating these competing claims empirically is the possibility that a regular pattern such as past-tense *-ed* is so high in frequency

* Fax: +1 604 822 9687.
E-mail address: stemberg@interchange.ubc.ca

0749-596X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2007.02.003

that it may be subject to ceiling effects: factors that have an effect on outputs of lower frequency (including irregulars) may not have an observable effect on regular patterns such as *-ed*, because performance is so good that it is close to ceiling.

An additional issue that is most often overlooked is the relation between morphology and general phonological processing: Are past-tense forms created in the same module in which basic phonological processing occurs, as is generally assumed in linguistic theories of the past two decades (e.g. Kiparsky, 1982; McCarthy, 2005)? Or are past-tense forms generated in a special module which is distinct from general phonological processing, as apparently (but tacitly) assumed by e.g. Rumelhart and McClelland (1986), Pinker and Prince (1988), and Pinker and Ullman (2002)? Some connectionist research emphasizes the possibility of phonological effects on morphology (e.g. Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002), but has been silent about whether these are the same phonological effects that are observable in the processing of all words. Stemberger (2002, e.g.) has argued that morphology is subject to the same basic effects as general phonological processing, so that they must occur in the same module. In this paper, I will provide further evidence that past-tense forms are generated in a way that makes them subject to the same effects of phonological processing that are observed on all words; this is true for both irregular and regular verbs. I will also argue that the superficially different behaviour of regulars and irregulars in one type of error does not depend on (ir)regularity per se, but instead can derive from phonological factors that are partially confounded with (ir)regularity.

One way to address ceiling effects is to focus on some subpart of past-tense usage in which a regular suffix such as *-ed* is used only rarely. If past-tense forms appear in the output on a relatively small proportion of trials, it is unlikely that ceiling effects will be present, and likely that other factors can express themselves. I will focus on the production of past-tense in complex syntactic environments containing the auxiliary verb *did* or *didn't*. Children produce past-tense forms in such environments (as a type of error known as *overtensing*) on about 15% of trials during an early stage of development. By focusing on an area where past-tense forms are common but nonetheless in a minority of trials, we may be able to detect effects with regular *-ed* that are less obvious in other circumstances. I will show that such errors, while superficially syntactic in nature, are subject to phonological and lexical effects, such that errors are more common if the resulting form benefits from output biases in the phonology or if the target word is of high frequency (even though the target form is the default output pattern). I will show that these output biases interact with the pronunciations of regular vs. irregular forms in such a way that regular forms are overall less likely to undergo overtensing errors. I discuss the ramifications of the results for models of language production, arguing for a particular view of language processing (the integrated multiple competitor approach) that can be incorporated into both symbolic and connectionist models.

Marking past-tense in English

In English, past-tense is mapped onto form in an inconsistent manner, as shown in Table 1 with the regular verb *sneeze* as illustration. There are two competing ways in which past-tense is mapped onto form. In one pattern, past-tense is mapped onto the verb, via the addition of *-ed* in regular verbs, and by various other changes in irregular verbs. In the other pattern, the verb appears in a nonfinite (uninflected base) form, while past-tense is expressed via the addition of the auxiliary verb *did* or *didn't*. Which form is used cannot be reduced entirely to whether the sentence is positive vs. negative, a statement vs. a question, or emphatic vs. nonemphatic. It should be noted that, as far as the verb itself is concerned, there are two patterns which are considered to be defaults: if there is no auxiliary, the default is to add *-ed* to the verb; if there is an auxiliary, the default is to use the nonfinite form of the verb (the uninflected base). The occurrence of both patterns under very similar conditions of meaning presumably adds an extra challenge in the acquisition and processing of English.

Andrews (1990), working within the theory of Lexical Functional Grammar, pointed out that a special mechanism is needed to keep the two patterns separate. In a simple system, there is a rule that expresses past-tense on the verb, and another rule that expresses past-tense via an auxiliary verb. There is no inherent reason why tense could not be expressed on both. Indeed, perfect aspect in English is expressed via an auxiliary verb *and* via affixation on the main verb: *he* ***has*** *sneeze****ed*** *every day so far*. In English, there is a special mechanism to enforce a special dependency, whereby past-tense is *not* mapped onto the main verb if it is mapped onto an auxiliary (e.g., *ate*, *didn't eat*, but **didn't ate*). Andrews suggests that this mechanism must be learned, and that errors can arise in which the dependency is erroneously ignored, leading to the double expression of past-tense in *overtensing* errors: **didn't br**o**ke* and **didn't s**a**ng* (instead of *didn't br**ea**k* and *didn't s**i**ng*). By adulthood, this dependency has definitely been learned. Even in speeded experimental tasks (e.g. Stemberger & Middleton, 2003), adult native speakers of English produce regular forms such as *sneezed* accurately in simple positive statements more than 99% of the time (only occasionally producing uninflected base forms, known as *zero-marking errors*, such as **sneeze*), and produce regular forms such as *didn't sneeze* accurately in simple negative