

Barely There: Hard-to-Detect Auxiliaries Shed Light on Children's Acquisition of French

Cristina D. Dye¹

1 Overview

The goal of the present paper is to shed new light on the status of functional elements, and in particular on the status of auxiliaries in first language acquisition. Although the status of function words in child language has been under investigation for quite some time now, it continues to be an intensely debated topic in the field. According to one view, which here I refer to as the early syntactic sensitivity view, syntactic competence for grammatical categories, such as auxiliaries, is available to toddlers and even infants. According to another view, which I refer to as the deprivationalist view, auxiliaries and other grammatical categories are absent early on, children's speech being considered "telegraphic" (e.g., Brown 1973).

The present study focuses on auxiliaries in child French and aims to evaluate these two views using a new corpus of child speech samples and more precise methodology. Throughout this paper, I will attempt to draw the reader's attention to the impact of research methodology on our understanding of child language acquisition. We will see that early language is less impoverished than previously thought, a fact with important consequences for future theories of acquisition.

2 The Early Syntactic Sensitivity View

The early syntactic sensitivity view has been supported by a number of experimental findings from comprehension studies, which indicate that from a very early age children are sensitive to function words. For example, infants and toddlers have been shown to be able to detect function words in the

¹I am grateful to Barbara Lust, Carol Rosen, Yasuhiro Shirai, and John Whitman, as well as Réjane Frick, Yumiko Nishi, Marc Brunelle, Andrew Spencer, Yarden Kedar, Allan Dye, and to the Electric Wilburland Sound Studio in Ithaca, NY. I also thank the daycares Les Petits Malins, Crèche de L'enfaneluche, Halte-garderie Porte de Villiers, their directors Christine Milano, Nicole Dambrin, Isabelle Fregnac, and the parents and children who participated. This research was supported by a Sicca Research Grant and a Cognitive Studies Fellowship from Cornell University.

speech stream in English (e.g. Shady 1996, Hirsh-Pasek 2000, Golinkoff, Hirsh-Pasek and Schweisguth 2001) and German (e.g. Höhle and Weissenborn 2003).

Additionally and more importantly, several studies have shown that infants and toddlers are able to distinguish between different types of function words (e.g. determiner vs. conjunction) and that they are capable of using their detection of function words in the speech stream to compute word reference and to determine the grammatical category of a given content word (e.g. Gerken and McIntosh 1993, Höhle, Weissenborn, Keifer, Schultz and Schmitz 2004, Kedar, Casasola and Lust 2006). For example, using a preferential looking task, a recent study by Kedar, Casasola and Lust (2006) tested American-English infants of 18 and 24 months on their ability to detect ungrammaticalities caused by manipulating a function word in simple (interrogative) sentences. Infants heard grammatical sentences in which the determiner ‘the’ preceded a target noun, and also three ungrammatical conditions in which ‘the’ was either absent, was replaced by a nonsense function word ‘el’, or was replaced by the conjunction ‘and’. Both the 18- and 24-month-old infants oriented faster and more accurately to a visual target following grammatical sentences. Kedar et al. (2006) argue that by 18 months of age, infants use their knowledge of determiners in sentence computation and in establishing reference.

3 The Deprivationalist View

According to the deprivationalist view, which is to a large extent based on corpus studies, function words are essentially lacking in children. More than three decades ago, based on the investigation of the production of function words in natural speech samples from the three Harvard children, Adam, Eve and Sarah, Brown (1973) concluded that child language is structurally reduced. Much subsequent research similarly concluded that grammatical words and grammar in general are absent or impoverished in children. For example, Radford (1990) posits that children lack functional categories and the associated functional projections (cf. also, e.g., Clahsen 1990, Bickerton 1999). Some studies, such as Tomasello (1992), go as far as to propose that children lack abstract grammar altogether. A study of four Dutch children specifically notes that “overt auxiliaries are *never* observed [in early language]” (Wijnen 1997:21, italics added).² A fairly recent study of child

²The Dutch study does not specify the ages at which the children started producing overt auxiliaries. It is implied that auxiliaries are absent at least in the first sessions, during which the ages of the children were 1;9, 1;11, 2;0 and 2;7, respectively.

French claims that children go through a “first stage [...] [with] no signs of instantiations of INFL,” that is, no auxiliaries or subject clitics. The three children concerned were aged 2;0, 2;2, and 2;3, respectively, during the said first stage (Schlyter 2003:23). (More detailed literature reviews for both the early syntactic sensitivity view and the deprivationalist view are found in Dye (2005:ch. 2) and Lust (2006:192-207).)

4 Limitations of Data Used in Previous Corpus Studies

However, the deprivationalist view is based on somewhat problematic evidence in that the data used in previous corpus-based studies suffers from critical limitations. A number of studies are based on small sample sizes and even single cases (e.g. Tomasello 1992), rendering the claimed findings less generalizable. At the same time, several studies rely on the same (few) subjects, for example, studies on English function words are often based on Brown's three subjects.

Furthermore, some studies are based on diary reports (e.g. Tomasello 1992). Unavailability of audio recordings prevents ambiguities arising in analyses from being resolved. One also needs to take into consideration the fact that even when there are recordings, much previous work has relied on dated analog audio recordings, or on video recordings only (in which sound quality is low). It is now known that function words in many languages, including English, have reduced prosodic salience, being prosodically deficient (e.g. Zec 2002), and thus are easy to overlook during transcription. A sociolinguist with extensive experience in studying auxiliaries in *adult* speech referred to function words as “small bits of sound that are often hard to hear” and said that “in fast speech they all but disappear” (Labov 1995:28–9). It is reasonable to expect that auxiliaries may be even harder to detect in child speech, and particularly so when only video recordings are available or when the audio recordings involve outdated technology.

Additionally, a number of studies use corpora that lack information concerning the linguistic and pragmatic context of children's utterances; without such information it is difficult if not impossible to establish whether or not a given utterance is grammatical. Limitations of the types noted above suggest an urgent need for improved methods.

5 A New Approach to Compiling Child Corpora: The Dye 2005 Corpus

To address the issues discussed in the previous section, I have compiled a new child French corpus, the Dye Corpus (described in more detail in Dye 2005). Given the object of my investigation—auxiliaries—it was crucial to have detailed data. The Dye Corpus consists of over 5,000 child utterances based on cross-sectional natural speech samples containing approximately 3,438 verb clauses; the high verb density provides for robust analyses. The participants were 18 normally-developing French monolingual children with ages ranging from 1;11 to 2;11 (mean age 2;5). The children were recruited through day-care centers in Paris and Nancy, France. They were interviewed individually at the day-cares in a quiet room. Average interview length was thirty minutes.

My primary goals in compiling the new corpus were a) to ensure the availability of discourse contexts for the linguistic structures under investigation, namely, verbs and auxiliaries, b) to ensure comparability among speech samples from different children, and c) to ensure the audio quality necessary to capture functional items. To ensure that children had ample opportunity to produce the targeted items, I selected a set of activities, toys and conversation topics that trigger utterances with verbs and auxiliaries. To render the speech samples obtained from different children more comparable, I attempted to standardize as much as possible several aspects of the interview process, namely, conversation topics, games, toys, interviewers, interviewer training, and interview location. To ensure that the recordings had the audio quality necessary to capture functional items, a new recording set-up was used (Figure 1). The new recording set-up consisted of a minidisc recorder, which attached to the interviewer's belt, and a set of microphones, which were worn in the ears, like headphones. This set-up offers the advantages of being mobile, allowing proximity to the child at all times, and of being very unobtrusive, thus not distracting or intimidating the child. To obtain high-fidelity samples, I used a digital Sharp IM-M T880 minidisc recorder, which made it possible to clearly capture children's voices even when they turned away or whispered. I used a Soundman OKM binaural stereo condenser microphone consisting of a pair of microphones designed to capture sonic information in a manner similar to the human ear, thus producing very realistic recordings. This model works on phantom power from the recorder, thus eliminating the noise associated with a powered microphone. The sessions were also videotaped.



Figure 1: The new recording set-up used in his study (arrows point to the microphone and recorder)

The recordings were digitally edited using *Cool Edit* to enhance the children's voices in relation to the background noise. The interviews were then transcribed with WAVpedal 5.0, transcription software that allows transcription directly from the computer, without going through analog, and thus avoiding loss of quality. Audio samples were transcribed by trained native speakers and checked by the author, all of whom were present at the interviews. Transcription was carried out in such a manner as to capture as closely as possible the actual phonological form heard on the recording while using French orthography. Where necessary, transcription and coding referred to spectrogram analyses carried out in *Praat*. Coding followed systematic procedures based on the Cornell University Virtual Linguistics Lab Research Methods Manual (Lust, Blume and Ogden, in prep.). Analyses of the speech samples were based on audio, video, and transcribed data. For further information on coding, please see Dye (2005:76–81).

6 Results

The results of this study revealed that children produce a variety of auxiliary/modal forms and types, as exemplified below. Additional examples of the utterance types presented in this section may be found in Dye (2005:chapter 3). For the sake of simplicity, the term *auxiliary* will be used as a cover term for both auxiliaries and modals.³ I will start by presenting

³Strings of the type modal + nonfinite e.g., *faut partir* 'must leave' are treated as monoclausal, as such collocations are very common in colloquial French, functioning

examples with full auxiliary forms and then move on to reduced auxiliary forms. An example of an utterance with a full auxiliary, here the auxiliary *peut* ‘can’ in its target form, is provided in (1) and belongs to the youngest child in the study.

(1) l’éléphant, i **peut** pus, i **peut** pus tomber (age 1;11)

‘the elephant, he **can** no-longer, he **can** no-longer fall’

Among the different types of auxiliaries, the most frequent are the past tense auxiliaries *avoir/être* ‘have/be’, followed by the immediate future auxiliary *aller* ‘gonna’, followed by *pouvoir* ‘can’, followed by *vouloir* ‘wanna’, followed by *falloir* ‘gotta’. Utterances with full auxiliaries are found in all children, from the earliest ages, as shown in Figure 2.

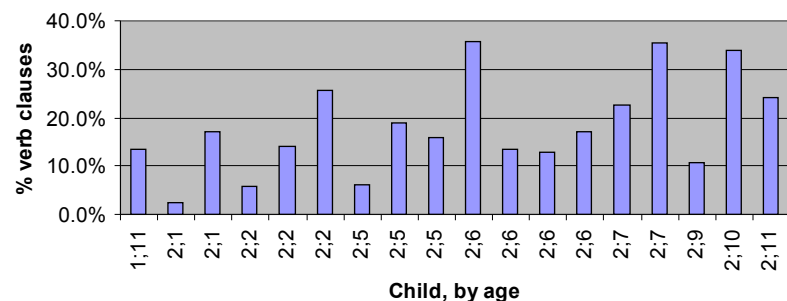


Figure 2: Full-aux clauses as a percentage of all verb clauses, by individual child

In addition to full auxiliaries, children also produce reduced auxiliary forms. These may be phonologically reduced auxiliaries, as in (2), or they may be phonetically reduced, as in (3). An example of a phonologically reduced auxiliary is the form *oudrais*, with initial *v*-deletion instead of the target *voudrais* ‘would-like-to’:

like aux + nonfinite, and other Romance varieties treat them as monoclausal. Schlyter (2003:21) makes the same point, based on the fact that French modals and auxes are used in very early stages, by both L1 and L2 learners, as markings of Tense/Mood/Agreement. We must recognize a divergence between what syntactic theory tells us, and what these strings say in terms of extreme frequency. Based on these functional criteria, I treat them as parallel to other aux + nonfinite structures (e.g. past tense), though I recognize that, for whatever reasons, they remain bi-clausal in French under syntactic tests.

(2) j'oudrais enlever les autocollants. (age 2;6)

'I would-like-to remove the stickers.'

Some auxiliary forms are barely audible; such forms are phonetically reduced compared to the rest of the utterance, having lower intensity. For example, in (3), the element immediately preceding the verb was initially barely audible.

(3) _lever ça (age 2;6)

'_ lift-INF this'

Spectrogram analyses of this utterance revealed the presence of a fricative, [f]/[v] (Figure 3). Inspection of the linguistic and pragmatic context of this utterance showed that in fact it occurs in a sequence where the child utters several times *faut lever* 'must lift' in self-addressed speech. The element preceding the verb in (3) is thus a partial pronunciation of the auxiliary *faut* [fo] as indicated in (3')

(3') f lever ça (age 2;6)

'must lift this'

Some of the children's auxiliary forms show more dramatic reduction. Among such forms, one may distinguish phonologically deleted auxiliaries, whose presence is signaled by preceding subject agreement markers, as well as monosyllabic "fillers." Given that in Colloquial French and particularly in child French, subject clitics are known to function as agreement markers,⁴ examples such as (4) where a subject clitic precedes a nonfinite verb, implicate the presence of a finite auxiliary (here possibly *va* 'gonna', based on the context).

(4) R⁵ et qu'est-ce qu'il va faire Timothée avec sa maman? (age 2;2)

'and what is it that he is-gonna do Timothy with his mom?'

C: i dormir

⁴See for example Ashby (1977), Auger (1993), Jakubowicz and Rigaut (2000).

⁵R = researcher, C = child.

‘he sleep-INF’

Examples (5) and (6) involve a monosyllabic element *é* [*e*], which I refer to as a “filler”, following Peters (1997), Pepinsky, Demuth and Roark (2001). It follows the subject agreement marker *z* ‘I’⁶ and precedes the nonfinite verb. Given its sentential position, the filler *é* [*e*] is a likely candidate for auxiliary status. In both (5) and (6), the context involves actions in the immediate future. In (5), the child has been asking to go play outside; the researcher asks what he intends to do outside and the child explains his intentions.

(5) R: Alors, qu’est-ce que tu vas faire là? (age 2;1)

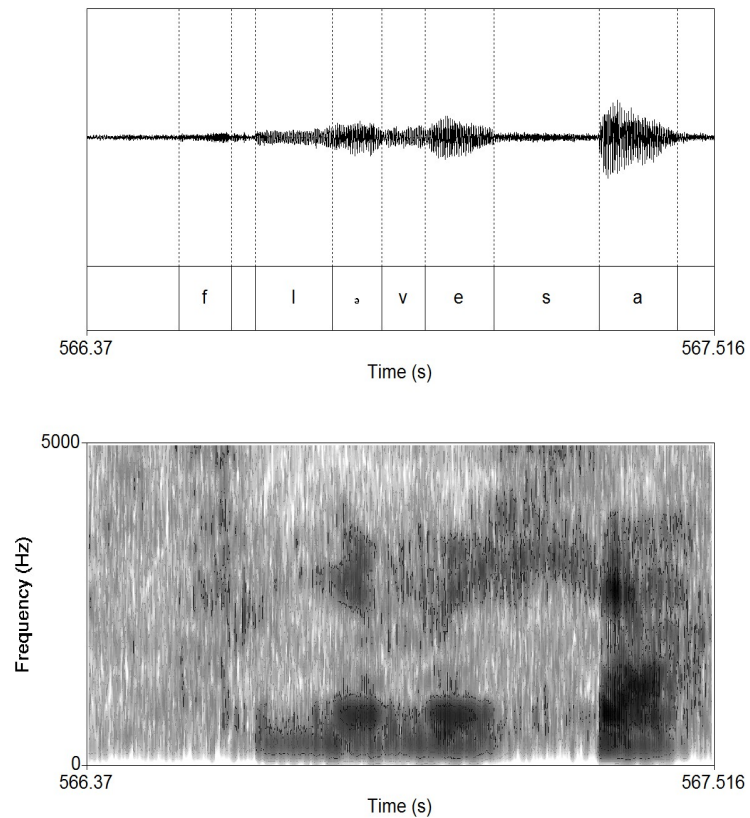
‘so what is it that you are-gonna do there?’

C: z’é [*e*] zouer

‘I é[*e*] play’ (‘I’m gonna play’)

⁶Z is the child’s version of the target subject clitic *j(e)* ‘I’.

F

figure 3: Spectrogram of *f lever ça*

In example (6), the child is asking the adult's opinion about placing a figurine in a certain position on the felt board.

- (6) C: z'ɛ̃[e] mett' celi-là, celi-là, là, à côté cé monsieur? (age 1;11)

'I ɛ̃ [e] put-INF that-one, that-one there next to this mister?' ('I'm gonna put...?')

R: uhum.

Both (5) and (6) refer to immediate future actions; *é [e]* is therefore likely to be a phonologically reduced form of the immediate future auxiliary *vais [ve]* ‘gonna/go’. Confirmation that *z’é* represents a reduction of *j’vais* is found in the speech of an older child where the form in question occurs as a main verb.⁷ In (7), it is clear that *é [e]* stands for *vais [ve]* ‘go’. In this example, the child is talking about her morning routine, which involves going to the day-care.

(7) C: j’prends l’déjeuner quand quand z’mè lève du lit (age 2;10)

‘I have the breakfast when when I *reflexive* get-up from-the bed’

R: oui, et puis après?

‘yes, and then, afterwards?’

C: z’*é* [e] à la crèche quand z’mè lève du lit

I *é* [e] to the daycare when I reflexive get-up from-the bed

‘I *go* to the daycare when I get up from bed’

Further support for fillers as phonologically reduced auxiliaries comes from a number of facts. First, it is well attested that auxiliaries are prone to phonological reduction and deletion, both diachronically and synchronically (e.g., Labov 1969, 1995). Second, in Modern French there are several auxiliaries with already reduced monosyllabic forms, e.g., *a* ‘has’ or *va* ‘gonna’. Third, and this represents a new discovery, the Dye corpus contains instances of auxiliary deletion in *adult* fast speech, as illustrated in example (8). Here the target *je vais te montrer* [ʒə ve t(ə) mɔ̃tʁe] ‘I am-gonna you show’ is realized as *je _t montrer* [ʒətmɔ̃tʁe], showing phonological deletion of the auxiliary *vais* ‘gonna’:

(8) ...regarde, je _t montrer (adult)

‘....watch, I _ you show-INF’ (‘I’m gonna show you ’)

⁷In French, the translation equivalents of the auxiliary *gonna* and of the main verb *go* are homophonous.

In Figure 4 below, I have added all the reduced auxiliary forms to the full auxiliary forms presented in Figure 2 above. Clauses with auxiliaries range from 5.1% to 38.8% of a child's total verb clauses. On average, 24.8% of children's verb clauses contain auxiliaries (see Appendix).



Figure 4: Clauses with full or reduced auxes as a percentage of all verb clauses, by individual child

To summarize the present results, auxiliaries and modals are evidenced in *all* children from the *earliest* ages, contra previous studies (e.g. Wijnen 1997, Schlyter 2003). It seems likely that the disparity between these results and those reported in earlier studies may be due to the improved methodology used here which made it possible to pick up monosyllabic weak elements (reduced auxiliaries, fillers, subject clitics) even when they were uttered fast or whispered. We have seen that the spontaneous productions of the eighteen French-speaking children studied here include a variety of auxiliary forms: full auxiliaries, phonologically reduced auxiliaries, barely audible auxiliaries, even more-reduced auxiliaries in the form of what I have referred to as “fillers”, and additional more-reduced/deleted auxiliaries, whose underlying presence is signaled by a preceding subject agreement marker.

7 Conclusion

The present study, carried out with improved methodology compared to previous corpus-based work, has uncovered new evidence of auxiliaries in young children's productions. This research strengthens previous experimental findings from discrimination and comprehension studies regarding

children's early knowledge of functional elements, by providing converging evidence from production. These results do not support the deprivationalist view of language acquisition. The present study additionally demonstrates the impact of research methodology on our understanding of child language acquisition and points to the need for new child corpora that will allow more precise examination of early language.

Appendix

Child age	Total verb clauses	All aux clauses	% All aux clauses
1;11	180	35	19.4%
2;1	196	10	5.1%
2;1	76	17	22.4%
2;2	106	16	15.1%
2;2	107	21	19.6%
2;2	278	89	32.0%
2;5	166	13	7.8%
2;5	305	59	19.3%
2;5	189	55	29.1%
2;6	188	73	38.8%
2;6	207	38	18.4%
2;6	109	19	17.4%
2;6	188	48	25.5%
2;7	97	27	27.8%
2;7	510	182	35.7%
2;9	93	14	15.1%
2;10	239	81	33.9%
2;11	204	54	26.5%
total	3438	851	24.8%

References

- Ashby, William J. 1977. *Clitic Inflection in French: An Historical Perspective*. Amsterdam: Rodopi.
- Auger, Julie. 1993. More evidence for verbal agreement-marking in colloquial French. *Linguistic Perspectives on the Romance Languages: Selected Papers from the 21st Linguistic Symposium on Romance Languages, Santa Barbara*,

- California, 21-24 February 1991, ed. W. J. Ashby, M. Mithun, G. Perissinotto, and E. Raposo. 177–198. Amsterdam, The Netherlands: John Benjamins.
- Bickerton, Derek. 1999. Creole languages, the language bioprogram hypothesis, and language acquisition. In *Handbook of Child Language Acquisition*, ed. W. C. Ritchie and T. K. Bhatia. 195–220. San Diego: Academic Press.
- Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge, Mass: Harvard UP.
- Clahsen, Harald. 1990. Constraints on parameter setting: a grammatical analysis of some acquisitional stages in German child language. *Language Acquisition* (1), 361–391.
- Dye, Cristina. 2005. *Identifying Auxiliaries in First Language Acquisition: Evidence from a New Child French Corpus*. Doctoral Dissertation, Cornell University.
- Gerken, LouAnn and B. McIntosh. 1993. Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29(2):448–457.
- Golinkoff, Roberta Michnick, Kathryn Hirsh-Pasek, and Melissa Schweisguth. 2001. A reappraisal of young children's knowledge of grammatical morphemes, vol. 1. In *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, ed. J. Weissenborn and B. Hohle. 167–188. Amsterdam: John Benjamins.
- Hirsh-Pasek, Katherine. 2000. Beyond Shipley, Smith, and Gleitman: Young children's comprehension of bound morphemes. In *Perception, Cognition, and Language: Essays in Honor of Henry and Lila Gleitman*, ed. B. Landau, J. Sabini, J. Jonides, and E. L. Newport. 191–208. Cambridge: MIT Press.
- Hohle, Barbara and Jurgen Weisseborn. 2003. German-learning infants' ability to detect unstressed closed-class elements in continuous speech. *Developmental Science*, 6:122–127.
- Hohle, Barbara, Jurgen Weisseborn, Dorothea Keifer, Antje Schultz, and Michaela Schmitz. 2004. Functional elements in infants' speech processing: The role of determiners in syntactic categorization of lexical elements. *Infancy* 5.3:341–353.
- Jakubowicz, Celia, and Catherine Rigaut. 2000. L'acquisition des clitiques nominatifs et des clitiques objets en français. *Canadian Journal of Linguistics*. 45:119–157.
- Kedar, Yarden, Marianella Casasola and Barbara Lust. 2006. Getting there faster: 18- and 24- month-old infants' use of function words to determine reference. *Child Development* 77.2:325–338.
- Labov, William. 1969. Contraction, deletion, and the inherent variability of the English copula. *Language* 45:715–762.
- Labov, William. 1995. The case of the missing copula: The interpretation of zeros in African-American English. In *An Invitation to Cognitive Sciences v.1: Language*, ed. L. R. Gleitman and M. Leiberan, 25–54. Cambridge, Mass.: MIT Press.
- Lust, Barbara. 2006. *Child Language: Acquisition and Growth*. Cambridge: Cambridge University Press.
- Lust, Barbara, Maria Blume, and Tina Ogden. in prep. *Cornell University Virtual Linguistics Lab Research Methods Manual: Scientific Methods for the Study of Language Acquisition*. Ithaca, NY: Cornell U.

- Pepinsky, Thomas, Katherine Demuth, and Brian Roark. 2001. The status of ‘filler syllables’ in children’s early speech. *Proceedings of BUCLD 25*, 575–586. Somerville, MA: Cascadilla Press.
- Peters, Ann M. 1997. Language typology, prosody and the acquisition of grammatical morphemes. In *The Crosslinguistic Study of Language Acquisition*, Vol. 5, ed. D.I. Slobin. 136–197. Hillsdale, N.J.: Lawrence Erlbaum Asst.
- Schlyter, Suzanne. 2003. Development of verb morphology and finiteness in children and adults acquiring French. In *Information Structure and the Dynamics of Language Acquisition*, ed. C. Dimroth and M. Starren. Amsterdam & Philadelphia: John Benjamins.
- Shady, M. 1996. *Infants’ Sensitivity to Function Morphemes*. Doctoral Dissertation. State, University of New York at Buffalo.
- Tomasello, Michael. 1992. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press.
- Zec, Draga. 2002. On the prosodic status of function words. *Working Papers of the Cornell Phonetics Laboratory 14*.

Department of Neuroscience
Georgetown University
Washington, D.C. 20057
cdd24@georgetown.edu