Published in final edited form as:

Theor Linguist. 2011 October ; 37(3-4): 143–152. doi:10.1515/thli.2011.010.

Sign language verb agreement and the ontology of morphosyntactic categories*

Mark Aronoff and Carol Padden

Inflectional morphology is of interest to syntacticians and morphologists alike because inflection provides a tangible window into syntactic categories. The very term morphosyntactic category denotes a category of a language that is identified on the basis of morphology that operates in the syntax. Where do such categories come from? We will call this the question of the ontology of morphosyntactic categories. Lillo-Martin and Meier (henceforth LM&M) raise three intriguing questions about the ontology of morphosyntactic categories: first, how do we know if a language exhibits a given morphosyntactic category; second, where do morphosyntactic categories come from; and finally, related to the second, why do languages display the morphosyntactic categories that they do?

Linguists have grappled with these questions for centuries. Many of the familiar morphosyntactic categories and terms, including tense, perfect, imperfect, person, number, gender, and case came to us from Classical Latin grammar (often as translations of Classical Greek grammatical terms), specifically from analysis of the morphology of Latin and Greek, since ancient Greek and Latin grammar was largely concerned with morphology. The categories and terms of these grammars were later adopted for use with other languages, as with the use of the terms perfect and imperfect for the Semitic suffixal and prefixal verb forms; sometimes they were adapted, as with the terms perfective and imperfective, coined by the pioneering Slavist Franc Miklošič, for the Slavic aspects, by analogy to perfect and imperfect.

No one thought much about the ontology of the categories that these terms implied until Franz Boas, the founder of cultural relativism in anthropology, famously declared that the traditional grammatical categories were too restrictive for the proper description of the native languages of the Americas. In a section bearing the title “Interpretation of Grammatical Categories” in his Introduction to *Handbook of American Indian Languages* (1911), Boas observed that “in a discussion of the characteristics of various languages different fundamental categories will be found.” What this meant in practical terms for Boas and his fellow field linguists was that one should not go into the field expecting to find only the grammatical categories that one had learned in school. The more extreme theoretical implication sometimes taken from Boas’s observation was that, in the now infamous words of Martin Joos (1957, p. 96), “languages can differ from each other without limit and in unpredictable ways.”

Chomskyan linguistics was in many ways an antithesis to Joos’s American Structuralist thesis, a framework that seeks invariance in human language as opposed to unlimited difference. Indeed, in the most recent versions of Chomsky’s theory, the so-called faculty of language in the narrow sense doesn’t vary at all, consisting solely of the operation Merge

*Thanks to Michael Cysouw and Stephen Wechsler for their helpful comments.

© Walter de Gruyter

Correspondence address: Stony Brook University, Mark.aronoff@stonybrook.edu.

(Chomsky 1995). Within the Chomskyan tradition, the most common approach to Boas's problem with respect to morphosyntactic categories has been modeled on the approach to phonology that generative linguistics inherited from the Prague school: universal distinctive features. The basic ontological solution for fundamental elements in generative phonological theory is quite simple: assume that there is a universal set of binary distinctive features, which we can discover by studying the phonological systems of a variety of languages. This universal set is assumed to be predetermined by what is usually called universal grammar or the human language faculty. Individual languages draw from this universal set in constructing their phonologies. This ontology has a variety of empirical weaknesses. For one, how do we handle the absence of a set of features in a given language? For example, although some languages make phonological use of tone, most do not. Does this mean that whatever distinctive features drive the tonal distinctions are absent in the grammars of speakers of nontonal languages, or only that they are dormant or perhaps, never having being activated, that they have somehow atrophied? A strict universalist would presumably want to argue for the latter, but it leaves us in the uncomfortable position of having to say that any feature that is distinctive in some language must be part of the universal set. Another serious weakness is that, although we assume that there is a universal set of phonological features, no one knows how to determine what the set of universal features is. One sign of this weakness is the fact that little effort has been spent in the last forty years on investigating the question of universal distinctive phonological features (although phonologists often assume that the features are somehow valid). Another question is how or why a given feature becomes active in a given language. For example, many linguists have explored the topic of *tonogenesis*, how distinctive tone arises. It turns out that there are many paths to the distinctiveness of tone, made more complex by the fact that there are many ways in which tone can be used systematically in languages.

The same type of ontological solution that has been used for phonological features can be applied to morphosyntactic categories, albeit with similar empirical difficulties (Haspelmath 2010). A strict universalist would say that the morphosyntactic categories and feature values of all the world's languages are all given in advance by the language faculty, although a given language will manifest only a subset of them. One linguistic theory that has made especial use of innate features is that of topographic syntax, which is concerned with functional categories involved in verb inflection. Shlonsky (2010) writes that: "familiar arguments from the poverty of stimulus militate in favor of the hypothesis that the functional features and their hierarchical arrangement are wired into the grammar and not learned or otherwise acquired through experience." This particular universalist framework commits us to believing that all morphosyntactic categories (maybe as many as 400 of them!) are not only given in advance but also somehow present in all languages, part of the language faculty, regardless of whether they are expressed in a particular language. The opposite point of view from universalism would hold, following Joos's remark to its logical conclusion, that every language has its own morphosyntactic categories. Where these categories come from in either sort of framework is not clear, but we will return to this question later.

Regardless of which view we hold about the set of possible categories, extreme universalism, extreme individualism, or something in between, all linguists agree that the basic problem of identifying what categories are *active* in a given language can be profitably approached through inflection: the morphosyntactic categories of a language structure its inflectional paradigm and must therefore be expressed in the morphology of the language. We can accordingly discover for any given language what its morphosyntactic categories are by looking at its morphology. This is what LM&M have done for ASL and a number of other signed languages.

LM&M argue that ASL and other signed languages have two distinct values for the category Person: first (or speaker) and non-first. In this regard, signed languages differ from spoken languages, which have three values: first (speaker), second (addressee), and third (neither speaker nor addressee). Their most powerful arguments for this person system being part of the grammar of these languages, rather than purely pragmatic, are indeed morphological and are based on Saussure's principle of the arbitrariness of linguistic signs: the fact that a sign is arbitrary shows that it is part of a linguistic system, though it is now clear from much work on sign languages that a sign can be part of a linguistic system even if it is highly iconic. Thus, some sign languages show arbitrary language-specific signs for person, but for first person only: in Japanese Sign Language, one sign for first person (used both for pronouns and for inflectional markers) makes contact with the speaker's nose; in Brazilian Sign Language, there is a specific possessor hand shape restricted to first person. The fact that the forms of these signs cannot be predicted on purely pragmatic grounds is strong evidence for the grammaticalization of the first person category in these languages: for each of these languages, because the form of the first person marker is to some extent arbitrary, it follows that it is listed in the lexicon of the language and must therefore be learned. This means that the category first person is part of the grammar of the language. Even more remarkable are idiosyncratic first person object agreement forms for certain verbs in ASL (e.g. for the verb CONVINCe, which has two possible forms, both idiosyncratic) and Danish Sign Language (e.g. for the verb COMFORT). Here, not just the person marker but the entire verb form must be learned.

It is worth noting that the argument for morphosyntactic categories from single forms is not new. The most remarkable instance that we know of is the following remark by Bernard Bloch about the English forms *am* and *are*, which are the only two verb forms in English that encode first person singular present and neither first singular nor third singular present, respectively: "This constitutes the most striking idiosyncrasy of the verb [BE]: that its inflection distinguishes categories not recognized by the morphology of any other verb" (Bloch 1947, p. 407). But it is precisely these two forms and these alone that force us to posit these two slots in the paradigms of verbs in English.

Just as striking but in the opposite direction, recall that there are no arbitrary forms for non-first person in the sign languages that LM&M discuss, no special second-person markers analogous to the Japanese nose sign and no irregular complete verb forms like those for the first person object form of CONVINCe. Furthermore, there are no idiosyncratic or irregular forms for subject agreement, regardless of person. Why should these observations be true? The answer to both these questions and also to the question of why signed languages have no second- vs. third-person distinction beyond that of first lies, we maintain in the visual world that sign language inhabit, specifically in the nature of pointing.

Let us consider in more detail the nature of the category person and its relation to pointing in space, what is usually called *deixis*, a word borrowed directly from Classical Greek, where it also means 'pointing'. Person is a form of deixis. Otto Jespersen (1922) coined the term *shifter* to refer to words "whose meaning differs according to the situation" (p. 123) and notes that "The most important class of shifters are the personal pronouns" (ibid.). As Cysouw (2003) remarks, "All linguistic elements that are used in a deictic ('pointing') function are shifters. Locational deixis (*here, there*), time deixis (*now, tomorrow*), and participant deixis (*I, you*) all involve shifting reference of linguistic elements" (p. 5). In all languages, spoken or signed, words like *here, there, this*, and *that* must be accompanied by a pointing gesture when they are used deictically in space. It is simply infelicitous in any language to respond to the question "where did they go?" with the word "there" without pointing towards the spot where they actually went, although a speaker or signer may not always use a hand to point but sometimes something more subtle, like eye gaze. Only the

word *here* can be used deictically without pointing, and in that case it means ‘where the speaker is located’, an important point to which we will return.

To repeat Jespersen’s observation, the most important class of shifters comprises the personal pronouns. What is most remarkable about first and second person pronouns in spoken languages is that their deixis/shifting/reference is built in and so does not require an act of pointing: the first person pronoun points inherently to the speaker, whoever that is, while the second person pronoun points inherently to the addressee, whoever that is. Thus, there is no need to actually point when using these forms deictically, though there is no prohibition against a speaker gesturing towards their own body or the addressee’s when using the first or second person form. Third person pronouns, though, cannot be used deictically without pointing to the particular location in space where the referent is located. This is, of course, also true in signed languages with both non-first person pronouns and agreement: their location varies according to the locus of the referent. In this regard, deictic words in spoken languages are just like what LM&M call “directionality”. Both speaker and signer must point to a particular locus when using these forms. To summarize, deixis ‘pointing’ is inherently visual and both the “directionality” system of signed languages and the shifters of spoken languages rely on the visual nature of deixis. Furthermore, in both spoken and signed deixis, the spot to which one points depends on the actual or assigned locus of the referent. Importantly, this locus will vary in signed deixis, except in one case, the locus of the signer or speaker. This locus is constant, because this locus is where the signing or speaking emanates from: the body of the signer/speaker.

It is also important to note that the body is not *exclusively* the speaker. The body can also be a body: a signer can deictically point to body parts, or sign in the direction or making contact with the body, to hold a gun to the temple, or chest, or anywhere else on the body, to brush hair, teeth, etc. The body ceases to be speaker, and becomes a very detailed physical object with multiple possible locations.

We are now in a position to understand why the first person form in sign languages is the only one that can vary arbitrarily from language to language and that can be lexically specified for an individual verb: it is the only person marker that can be fixed in sign languages; all the others are deictic by their very nature and literally shift, not just in their referent but also in their location in space, every time they are used. They thus have no opportunity to acquire a fixed form. At best, a verb may be lexically specified for a certain height for *all* objects, but that still allows the deixis to shift in the horizontal plane, as Lillo-Martin and Meier emphasize.

But why can the form of the first-person object marker become arbitrary or lexically specified, but not the form of the first-person subject marker? Again, part of the answer lies in the variability of the form of the agreeing verb that is induced by directionality. As Lillo-Martin and Meier discuss at some length, subject agreement is optional in sign languages, while object agreement is obligatory. This has consequences for the form of the verb: since the form of the verb does not always have to agree in directionality with the subject, there is no opportunity for the first-person form of the verb to become fixed, a necessary condition for lexicalization. At a deeper level, it has been argued elsewhere (Meir et al. 2007) that the body of the signer serves another iconic function besides marking the speaker; it marks the subject. This can be seen in verbs that do not agree, what are usually called ‘plain’ verbs. With these verbs, which comprise the majority of verbs in all sign languages, the hands mark the predicate, while the body marks the subject. There are also agreeing verbs (e.g. ASL ANSWER) which cannot agree with the subject and this is because the part of the verb sign that could vary with the locus of the subject is already fixed on a particular place on the body because of the preeminence of ‘body as subject’. We see, then, that the inflectional

forms of sign language verbs argue for their having a system with one feature within the category person with two values, first and non-first, and that this system makes sense in terms of the modality in which sign languages are encoded: the visual modality leads sign languages to have this particular person system and even to have arbitrary markers for first person but not for non-first person.¹

Let us turn briefly to the category of number. LM&M say that agreement verbs in sign languages vary in both person and number, but in fact their arguments for number agreement are weaker. As they observe themselves, “nonfirst plural forms are compositional.” They do identify a number of arbitrary morphological forms in signs for first-person plurals similar to those that they use to motivate the special status of first-person singular, but they skirt a fundamental problem, discussed at length by Cysouw 2003, and Wechsler 2010, which is that “first- and second-person plurals always have ASSOCIATIVE semantics rather than REGULAR PLURAL semantics” (Wechsler 2010, p. 332). Therefore, the term first-person plural is a misnomer, since the category in question never refers to the plural of first-person but rather to the first person plus someone else. Consequently, first-person plural inflectional morphology cannot be used as an argument for the morphosyntactic reality of the category plural in sign languages.

What light does all this shed on Boas’s question? In ‘deciding’ to realize certain syntactic categories, features, and values morphologically, does a language select from a list given in advance by a higher authority or do the categories, features, and values emerge from other properties of the language situation? And what drives a language to select certain features over others?

In the case of person, Wechsler has made the very strong claim that the two privative features [*speaker*] and [*addressee*] driving the distinctions among types of pronouns and agreement involve self-ascription and are closely intertwined with human psychology, especially theory of mind. Wechsler specifically criticizes a Universal Grammar approach that involves binary features that are somehow specific to human language rather than emerging from wider cognitive abilities like theory of mind (Wechsler 2010, p. 341). Wechsler emphasizes the claim that all spoken languages include both these features in their grammars, which makes sense if the features are embedded in human cognition. What, then, about sign languages? Does the fact that the feature [*addressee*] does not function in these languages mean that, in some Whorfian sense, signers do not distinguish addressees from other participants in the discourse? That must be wrong. Instead, what it suggests is that the distinction between addressee and others, though available to signers on account of their being normal humans, is not obligatorily made in the grammar of their language, in the same way that a speaker of a spoken language without obligatory tense inflection like Vietnamese is not presumed to live in a timeless world; such speakers are simply not obliged to mark the time of an event with respect to the reference time of an utterance. Similarly, signers are not obliged to mark the distinction between addressee and other non-first person but that does not mean that signers do not understand the self-ascriptive nature of second-person reference.

But if addressee self-ascription of second person is so fundamental, why don’t sign languages mark it morphologically? Here the great influence of modality reemerges: pointing is fundamentally visual, which means that pointing is always incorporated into the linguistic system of sign languages in a way that is completely inconceivable for spoken

¹Wechsler (2010) claims that person features in spoken languages are privative and that there are two, speaker and addressee. This claim is embedded in a much larger discussion of the semantics of person in terms of self-ascription. As far as we can tell, an account along the lines of Wechsler’s would work for sign languages.

languages. Indeed all signers, even the least linguistically organized home signers (Goldin-Meadow 2003), use pointing indexically. But this very indexicality and the resulting variable form of all but the first-person sign actually prevents the emergence of a secondperson sign with constant form, which is why we find the person system that LM&M describe for most sign languages.

References

1. Bloch, Bernard. English verb inflection. *Language*. 1947; 23:399–418.
2. Boas, Franz. *Handbook of American Indian Languages*. Vol. Volume 1. Washington, DC: Bureau of American Ethnology; 1911. Introduction; p. 5-83.
3. Chomsky, Noam. *The Minimalist Program*. Cambridge, MA: MIT Press; 1995.
4. Cysouw, Michael. *The Paradigmatic Structure of Person Marking*. Oxford: Oxford University Press; 2003.
5. Goldin-Meadow, Susan. *The Resilience of Language: What Gesture Creation in Deaf Children can tell us about How All Children Learn Language*. New York: Psychology Press; 2003.
6. Haspelmath, Martin. Comparative concepts and descriptive categories in crosslinguistic studies. *Language*. 2010; 86:663–687.
7. Jespersen, Otto. *Language: Its Nature, Development, and Origin*. London: Allen and Unwin; 1922.
8. Joos, Martin. *Readings in Linguistics*. Chicago: The University of Chicago Press; 1957.
9. Meir, Irit; Padden, Carol; Aronoff, Mark; Sandler, Wendy. Body as subject. *Journal of Linguistics*. 2007; 43:531–563. [PubMed: 23066169]
10. Shlonsky, Ur. The cartographic enterprise in syntax. *Language and Linguistics Compass*. 2010; 4:417–429.
11. Wechsler, Stephen. What ‘you’ and ‘I’ mean to each other: Person indexicals, self-ascription, and theory of mind. *Language*. 2010; 86:332–365.