Quantitative measures of subjectification: A variationist study of Spanish *salir(se)**

JESSI ELANA AARON and RENA TORRES CACOULLOS

Abstract

By confronting variable use, the variationist method can reveal patterns of subjectification of grammatical morphemes. Applying this method to the analysis of salir(se) *'go out' variation in Mexican Spanish oral data, we conclude that subjectification is manifested structurally in the tendency for middle-marked* salirse *to co-occur with first-person singular or referents close to the speaker, positive polarity and the past tense. Further comparative dialectal and diachronic data indicate the origins of the* se-*marked form in physical spatial deviation. Usage of the form then extends to situations that denote deviation from social norms. We thus propose that the locus of subjectification of this counter-expectation marker is an increasingly speaker-based construal of expectation. This semantic change appears to proceed via absorption of contextual meaning in the frequently occurring +* de *'from' construction.*

Keywords: subjectification; variationist method; Spanish middle se; *counter-expectation.*

1.   Introduction: Subjectification and grammatical structure

Elizabeth Traugott's theory of subjectification sets forth the strong claim that the major type of semantic change is the development of explicit markers of subjectivity (Traugott 1989, 1995, 1999; Traugott and Dasher 2002). *Subjectivity* represents "a speaker's . . . point of view in discourse—what has been called a speaker's imprint" (Finegan 1995: 1). In the diachronic process of *subjectification*, meanings grounded in external objective reference change toward meanings based in the speaker's internal belief or attitude. Although linguistic analysis has largely privileged the "referential" over the "emotive" function of language (Jakobson 1960: 353–354), the subjectivity of language is pervasive (Benveniste 1966) and

subjectification has been a fashionable topic since the 1980s (cf. Stein and Wright 1995).

It is recognized that only certain linguistic elements constitute explicit expressions of subjectivity and that there are varying degrees of subjectivity (Traugott 1999: 179), yet operational measures of subjectification are still lacking (Torres Cacoullos and Schwenter 2005). In particular, we are faced with the question of what the structural correlates of subjectification are. We seem closer to an answer in the research on discourse markers and connectives, where subjectification involves the fixedness of internal structure together with the relaxation of syntactic constraints (Traugott 2003: 636–642). Thus, for example, the evolution of *instead of* from a locative nominal complex to a connective marking counter-expectation is manifested in syntactic generalization to more classes of nouns, *-ing* complements, and finally finite clauses (Schwenter and Traugott 1995). Syntactic correlates are explored by Company (to appear), who proposes that the subjectification of discourse markers involves ''syntactic cancellation'', as illustrated by Spanish pragmatic formulas originating in verbs. For example, the ditransitive valency of *dar* 'give' is annulled in the discourse formula *dale*, which indicates annoyance with obduracy.

But beyond discourse markers and connectives, subjectification also applies to productive grammatical(izing) morphemes and constructions. For example, in Langacker's (1990: 23) view, the development of the 'go' future involves a shift from intention of the subject to prediction by the speaker, while in Traugott's (1995: 36, 50) formulation, there is ''pragmatic reanalysis'' in the inferential (re)alignment of the subject-experiencer of *be going to* and the speaker's perspective. In this case, subjectification may be evinced by inanimate subjects or, more generally, by attenuation of syntactic subject agentivity (cf. Verhagen 1995: 107–108). Since here syntactic cancellation (Company, to appear) is not applicable, the question remains: Which structural elements can measure the degree of subjectivity in grammatical morphemes, and how?

Variable middle-marking of Spanish motion verb *salir* 'exit, go out, leave' offers an apposite site for investigating the structural correlates of subjectification in grammatical morphemes. Intransitive motion verbs are sometimes used with a reflexive pronoun clitic in Spanish, as in (1a). Here the verb *salir* appears with *se*, the third-person reflexive marker. The unmarked, or zero-form, counterpart is shown in (1b). The meaning difference attributed to the formal difference is one of unexpectedness, or going against the natural order of events—the norm is for people to exit from doors, not windows (e.g., Butt and Benjamin 2000: 362). Counter-expectation is an example of linguistic subjectivity since it involves the speaker's point of view (Traugott 1999: 179).

(1) a. "Se salió *por la ventana*". ¡*Se brincaba la ventana!*
          '"He *went out-SE* through the window." He would jump
          through the window!'   (Habla popular, 411)
     b. *en esos momentos venía* saliendo *en la puerta*
          'just then he was *coming out-0* at the door'

                                                   (Habla popular, 70)

In this study, we devise measures of subjectification and show how sub-
jectification of a grammatical morpheme is evidenced in distribution pat-
terns. These emerge from the quantitative analysis of variation between
marked and unmarked forms. To our knowledge, this is the first time the
variationist method (e.g., Labov 1969; Sankoff 1988a) has been applied to
the study of subjectification. The variationist viewpoint on language is
preoccupied with grammatical structure in discourse and the asymmetri-
cal relations between linguistic form and function therein (Sankoff 1988a:
141). Since the attribution of semantic motivations to speakers by the
analyst may be an *a posteriori* artifact of theoretical bias, the only access
we have to speakers' intentions in the choices of different forms is through
their naturally occurring use of language (Poplack and Tagliamonte 1999:
321–322; Sankoff 1988a: 154; cf. Dubois 1987: 811–812).

    Corpus-based quantitative case studies have begun to identify "struc-
tural patterns of subjectivity" (Scheibman 2002). The growing apprecia-
tion of the inseparability of linguistic usage and structure (e.g., Barlow
and Kemmer 2000; Bybee 2005; Bybee and Hopper 2001) requires such
an empirical approach, both because it can demonstrate subjectification
scientifically (cf. Diver 1985) and because it reveals patterns of speakers'
choices in discourse that are largely inaccessible to introspection. In the
variationist view, the recurrent patterns that constitute grammatical struc-
ture are reflected in frequencies of (co-)occurrence. A primary analytical
tool is "variable-rule" analysis, which models regularities in large bodies
of data by assessing statistical significance of contextual factors as well as
their magnitude and direction of effect (Sankoff 1988b).

    In what follows we will show that, synchronically, middle-marked *sal-
irse* is an explicit expression of subjectivity as evidenced in its probabil-
istic tendency to co-occur with the first-person singular or third-person
referents close to the speaker, positive polarity, and past tenses. Dia-
chronically, the subjectification of this counter-expectation marker ap-
pears to proceed via the absorption of contextual meaning in the fre-
quently occurring construction with *de* 'from'. Comparative dialectal and
diachronic data suggest that subjectification of *salirse* lies not so much in
the (increased) expression of counter-expectation per se, but rather in the
increasingly speaker-based construal of (counter-)expectation.

2.   *Salir + se*: An expression of counter-expectation

In modern Mexican Spanish data, we have found that *salirse* 'exit, go out, leave' occurs in the following four kinds of context: (i) leaving against obstacles or rules, (ii) leaving abruptly, (iii) leaving permanently, and (iv) leaving in order to have a good time. Let us look at some examples, drawn from the *Habla culta* and *Habla popular* Mexico City corpora (for a list of corpora used in this article see Appendix).[1] In (2a), the *se*-marked form is not a normal 'leaving school' when most students come out at the end of the day, but a violation of the rules, an escape. This is an event that evidently occurs against the normal expectations or desires of the speaker (the mother). Indeed, we find several examples of *salirse* in the context of exiting against obstacles or rules. In (2b), a woman goes out to meet a man secretly (*te ando tapando* 'I'm covering up for you'), and this exiting entails overcoming a fence, as indicated by *¿por donde?* 'from where?' and *¿cómo?* 'how?'. In (2c), the exiting is explicitly qualified as a *travesura* 'naughtiness', and has to be effectuated surreptitiously and rapidly (*corriendo* 'running').

(2)   *Salir*se 'exit against obstacles or rules, surreptitiously'
   a.   *Así es qu'el más grande, por burrito ... tampoco no ... ¡Se salía de la escuela, señora!*
        'So the oldest, the little fool ... didn't either ... He would *leave-SE* school, ma'am!'   (Habla popular, 411)
   b.   "*Dice mi tía que yo te ando tapando, te ando alcahueteando; pa que lo diga mi tía con provecho – dice –, vete con él.*" "*No, pos yo, ¿por dónde me salgo?*" ... *pos ... ai estaba todo cercado.* "*Yo ¿ ... cómo me salgo?*"
        '"My aunt says I'm covering up for you, I'm matchmaking for you; to make my aunt's word good, she says, go with him." "No, well, <u>from where</u> do I *go out-SE*?" ... well ... <u>everything there was fenced</u> ... "<u>how</u> do I *get out-SE*?"'
                                                            (Habla popular, 206)
   c.   *era una <u>travesura</u>: A la hora que todos estaban distraídos nosotros tratábamos de* salirnos <u>corriendo</u> *porque había cerca ... una señora muy limpia que vendía ... arroz con leche y natas*
        '[it] was a piece of <u>naughtiness</u>: When everybody was distracted we would try to *go out-SE* <u>running</u> because there was close by ... a very clean woman who sold ... rice with milk and cream'
                                                            (Habla culta, 122)

Another context of the *se*-marked form is abrupt exiting, as in (3a/b) *salirse volando/volada* 'fly out'.

(3) *Salir*se 'go out abruptly'

    a. *entró a la capilla, y vio al padre levantado del suelo como veinte centímetros, orando. Que* se salió <u>*volando*</u> *¿verdá?, y les fue a contar a los padres lo que había visto . . .*

        'he entered the chapel, and saw the priest lifted off the ground about twenty centimeters, praying. So he *went out-SE* <u>flying</u>, right?, and he went to tell the priests what he had seen . . .'

                                (Habla culta, 150)

    b. *ya me andaba mordiendo una tortuga un pie, -y que ¡me salgo* <u>*volada*</u>*! Ya no me volví a meter al agua*

        'a turtle was already biting my foot, so I *go out-SE* <u>flying</u>! I didn't go back into the water again'   (Habla popular, 58)

Salirse is also used when someone is permanently leaving a group, organization, or institution, as in (4a), where some fellows left the team (*y ahora ya empezamos con otros* 'and now we started up with others'). Similarly, the speaker uses *salirse* for quitting a job, as in (4b) (*y entré aquí* 'and I started working here'), or when leaving home for good, as in (4c) (*me separé de mi madre* 'I left my mother'). Indeed, Silva-Corvalán (1994: 123) characterizes the permanently leaving context illustrated by *Se salió del equipo* 'He (SE) left the team' as an "obligatory reflexive".

(4) *Salir*se 'leave permanently (a group, organization, job, home)'

    a. *Es que teníamos buen equipo, ¿verdá?, . . . pero* se salieron *unos muchachos y . . . y ora ya . . .* <u>*empezamos con otros*</u>

        'It's that we had a good team, right?, . . . but some fellows *left-SE* and . . . and now . . . <u>we started up with others</u>'

                                (Habla popular, 18)

    b. *y entré . . . este . . . a un molino. Ai duré cerca de unos siete años.* Me salí *luego,* <u>*y entré aquí*</u>

        'and I started working . . . uhm . . . in a mill. There I lasted about seven years. I *left-SE* after that, <u>and I started working here</u>'

                                (Habla popular, 439)

    c. *Enc. – ¿Y desde cuándo* se salió *de su casa, o . . . ?*

        Inf. – Tengo como aproximadamente unos doce años que yo <u>*me separé de mi madre*</u>*.*

        'Int. – And when did you *leave-SE* your house, or . . . ?

        Part. – . . . It's been about twelve years since <u>I left my mother</u>'

                                (Habla popular, 78)

Finally, *salirse* is used for going out, especially to have a good time, as in (5), where the speaker goes out *a la calle* 'onto the street' (5a), *por ai* 'just around' (5b), or *con las amigas* 'with the girlfriends' (5c). Going out

without a utilitarian purpose other than pleasure or just because one feels like it could be frowned upon and viewed as contrary to socio-cultural norms, especially when the one going out is a woman (Aaron 2004).

(5)   *Salir*se 'go out to have a good time'
    a.   *Me aburro; me salgo <u>a la calle un rato</u>, ¿no?*
        'I get bored; I *go out-SE* <u>onto the street</u> for a while, no?'
                                      (Habla popular, 137)
    b.   *Me escondía yo de mis padres,* me salía *yo <u>por ai</u>.*
        'I would hide from my parents, I would *go out-SE* <u>just around</u>.'
                                        (Habla popular, 85)
    c.   *Y luego, si* me salía *yo <u>con las amigas</u>, se enojaba. Me regañaba: "No; no debes de salir. Tú debes estar en tu casa".*
        'And then, if I would *go out-SE* <u>with the girlfriends</u>, he would get mad. He would scold me: "No; you shouldn't go out. You should be in your house"'.   (Habla popular, 59)

In all of these uses—exiting against obstacles, abruptly, permanently, or to have a good time—the *se*-marked form occurs where the speakers may be said to be not merely stating propositions but expressing their point of view.

   In an intriguing analysis, Maldonado (1999) characterizes Spanish *salirse* and other intransitive motion verb-plus-REFL-marker forms as "energetic" constructions, contrasting with "absolute" constructions in cognitive grammar terms (Langacker 1991: 389). Maldonado's energetic constructions analysis builds on Kemmer's (1993) proposal that the middle voice is characterized by a low level of distinguishability among participants. For example, whereas reflexive *me pregunto* (*a mí mismo*) 'I ask myself' distinguishes two, albeit co-referential, participants, middle *me pregunto* 'I wonder' blurs the distinction (Maldonado 1999: 20–21). In this analysis, a consequence of the low level of distinguishability among participants and hence the narrowing of the scope of predication is a focusing function, such that the marker *se* focuses on the change of state of the experiencer of the motion event, that is, *se* profiles a particular point in space rather than the entire trajectory of motion (Maldonado 1999: 353–399; cf. Kemmer 1993: 157). Through this focusing function, intransitive "absolute" construals gain dynamicity and become "energetic". Thus, Maldonado (1999: 369) accounts for dynamic effects of a rapid and abrupt situation, as in example (3), as a natural extension of the basic focusing strategy. He explains the 'going-against-normal-expectations' meaning of *se*-marked motion verbs as the last stage of development of energetic constructions: absolute > focused > unexpected (Maldonado 1999: 390).

Counter-expectation markers express a contrast "with what the speaker considers to be the norm" (Heine et al. 1991: 192; cf. Traugott 1999: 178). Accordingly, we may regard the development of an 'unexpectedness' meaning in motion verb-plus-REFL-marker forms as an instance of subjectification (or "pragmaticalization"; Maldonado 1999: 393) of a grammatical morpheme.

3.   Problem: Inherent variability (form-function asymmetry)

A hallmark of spoken language, however, is inherent variability (Labov 1969, 1972). Although it may be disconcerting to linguistic analysis, it is not always the case that language use reserves one form for one meaning. The pair of examples in (6) illustrates that the *se*-marked form (6a) and its unmarked counterpart (6b) may appear in nearly identical contexts, with apparently the same pragmatic function or semantic value: 'permanently leave home'/'against one's wishes'. These examples are from the same speaker, 10 lines apart.

(6)   *Salir*se-*salir* variation
    a.   *Pero cuando s'iba a casar el otro cuñado, le pidieron la casa a m'hija; entons, pus* nos *tuvimos que* salir.
       'But when the other brother-in-law was going to get married, they asked my daughter for the house; so, well <u>we had to</u> *leave-SE*.'   (Habla popular, 415, line 12)
    b.   *entons, ahora que ya m'hija ya se fue a vivir por allá y le pidieron la casa, le digo: "Pus <u>tenemos que</u> salir"*
       'so, now that my daughter went to live over there and they asked her for the house, I tell her: "Well <u>we have to</u> *leave-0*"'
                                 (Habla popular, 415, line 22)

In the next set of examples, the unmarked form *salir* occurs in contexts very similar to those we previously identified as characteristic of *salirse*: exiting abruptly (*volando* 'flying') (7), leaving permanently an organization or job (*no hay límite [de edad]* 'there's no age limit', *estoy pensionado* 'I'm retired') (8), and going out to have a good time (*con amigas* 'with girlfriends') (9).[2]

(7)   *Salir* 'go out abruptly'
    Cuando están en el restorán, acuérdate cómo … cómo sale *aquél <u>volando</u>, porque él ya … ya lleva como una obsesión …*
    'When they're in the restaurant, remember how … how that guy *goes out-0* <u>flying</u>, because he already … he already has like an obsession …'   (Habla culta, 416)

(8)   *Salir* 'leave permanently (an organization, job)'
  a.   *creo que son ahora nueve años como mínimo; para entrar. Pero,*
        para salir*, no hay límite; ... pueden durar todo el tiempo que*
        quieran en la Asociación
        'I think that now it's nine years [age] as a minimum; to enter.
        But, to *leave-0*, <u>there is no limit</u>; ... they can stay <u>in the Associ-</u>
        <u>ation</u> for as long as they want'   (Habla culta, 438)
  b.   *ya en mil novecient's sesenta,* salí ... *ya por edad, y estoy*
        pensionado
        'in nineteen-sixty, I *left-0* ... because of my age, and <u>I'm retired</u>'
                                                              (Habla popular, 158)

(9)   *Salir* 'go out to have a good time'
        me contaba mi mamá los minutos que llegaba yo a la casa. Y salir
        sola <u>con amigas</u>, como ahora, que se van al café ... pues no
        'my mom would count the minutes I took to get home. And *going*
        out-0 alone <u>with girlfriends</u>, like now, going for coffee ... no way'
                                                              (Habla culta, 296)

The set of examples (6)–(9) suggests that the two forms coincide in a range
of uses, whether as inferences from the context or as conventionalized
polysemies. Previous work on these forms, however, which has relied on
analysts' interpretations of handpicked examples, provides few clues into
how to measure subjectivity. Given variation, how can we empirically es-
tablish that the *se*-marked form is more subjective than the unmarked one?
It is futile, in our view, to attempt to determine relative degree of subjectiv-
ity for every individual pair of examples. One analyst's judgments are
susceptible to another's disputing, with no scientific advances achieved.

   Most apt for studying the distribution of the two forms is the variation-
ist method, which confronts the problem of form-function polyvalence
with the hypothesis of ''neutralization-in-discourse'': While two forms
almost always have contexts in which they have different meanings, we
do not expect the full panoply of distinctions to be pertinent each and
every time one of the alternate forms is used (Sankoff 1988a: 153–154).
Through systematic quantitative analysis of repeated occurrences of ap-
parently random alternations, the variationist method enables researchers
to discern patterns of co-occurrence with contextual elements. These pat-
terns show the structure of variable *salir se*-marking.

4.   Solution: Factors in variationist analysis as measures of subjectification

A total of 557 tokens of marked and unmarked *salir(se)* forms were ex-
tracted from the Mexico City corpora (not counting a handful of cases of

impersonal *se*, which precludes middle *se*, or cases with insufficient context [8 cases in all]). In these data, *salir* comprises 88% (491/557) of the tokens, while *se*-marked *salirse* makes up 12% (66/557).

Prior to the quantitative analysis, a crucial qualitative interpretative component of variationist methodology is the circumscription of the variable context. The variable context, or locus of variability in discourse, is circumscribed in consonance with the Labovian principle of accountability, which requires that every token be counted, not only those that lend support to the analyst's theoretical position. In particular, the principle of accountability specifies that analyses must account for ''every case where the variable element occurs in the relevant environments as we have defined them'' (Labov 1972: 72; cf. Milroy and Gordon 2003: 180–183; Poplack and Tagliamonte 2001: 89–91). Thus, we exclude from the quantitative analysis contexts where variation cannot occur.

Examination of all *salir(se)* occurrences led us to identify some uses that are exclusive to *salir*. These are copular 'turn out' and 'cost' uses, as in *va a salir tartamudeado* 'going to end up with a stutter' or *nos sale gratis* 'it's free', which make up 21% (103/491) of *salir* tokens. Other metaphoric uses, mainly 'appear', as in *salía en la televisión* 'he'd appear on television', and 'graduate', as in *ella había salido ya de la Universidad* 'she had already graduated from the University', comprise another 23% (115/491). Conventionalized expressions with *salir*, such as *salir con* 'say', *salir a* 'take after', *salir adelante* 'progress', and the discourse marker *sale* 'ok', make up 4% (18/491). Also apparently exclusive to the unmarked form, at least in the present data, is 'leave for a routine activity', such as leaving work to return home at the end of the day or leaving home to go to work, as in *salgo de aquí a las seis y media* 'I leave here at six thirty', and traveling from a geographic area, as in *para salir de la colonia, qué problema en la mañana* 'getting out of the neighborhood, what a problem in the morning'. These invariant contexts, totaling 51% (283/557) of the original data, were thenceforth excluded from the statistical analysis of *salir(se)* variation. All remaining 274 tokens were coded for a number of contextual features.

Our general working hypothesis is that, if *salirse* is more subjective than *salir*, it will appear at a greater rate in contexts that display subjectivity. More specifically, in ascertaining global tendencies in the data, if the *se*-marked form is more subjective, it should tend—in the aggregate—to co-occur with certain contextual elements. Contexts of occurrence are decomposed into a configuration of independent conditioning factors, whose contribution to speaker choice is modeled probabilistically in multivariate statistical analysis. With these factors we operationalize hypotheses about the choice between *salirse* and *salir*, extrapolated from

analyses of the Spanish middle voice and the more general subjectification literature, as well as from our own observations of the data. Each factor represents an operational measure of the subjectification of *salirse*.

We consider six environmental factor groups, which we hypothesize to influence *salirse* variation. The first three factor groups, co-occurrence of dative pronouns, grammatical person (first singular vs. others), and relationship to speaker (close vs. distant), endeavor to measure speaker involvement. First, we take co-occurrence of dative pronouns, as in (10), to provide a measure of affectedness, based on Maldonado's (1999: 394) argument that the involvement of the conceptualizer in counter-expectation constructions is shown by "the fact that dative *le* makes the use of *se* obligatory". Though ultimately left out of the multivariate analysis because of the low number of tokens (Milroy and Gordon 2003: 164), the rate of *salirse* with co-occurring datives conforms to the hypothesis and Maldonado's insight. Though not quite obligatory, *se* occurs with 66% (10/15) of the dative tokens.

(10)   *se lastimó . . . Un hueso* se *le* salió *aquí, del hombro*
       'he was injured . . . His bone *came out-SE* here [on him], from his
       shoulder'   (Habla popular, 204)

The second and third factor groups take into account features of the syntactic subject. In the grammatical person factor group, we expect the first-person singular to favor the *se*-marked form since it is probably the most acknowledged marker of subjectivity (Benveniste 1966; Scheibman 2002: 167). First (and second) person pronouns, as all deictics, "exhibit subjectivity" since their meanings are grounded in the speaker's point of view (Traugott and Dasher 2002: 22 and references therein). The "relationship to speaker" factor group that we introduce bears more justification. For third-person subjects, we hypothesize that referents close to the speaker should favor the *se*-marked form. We reason that a "speaker's imprint" (Finegan 1995: 1) will be imparted more when the speaker has emotional ties to the people talked about, as in (11a), where the speakers are talking about sons. Referents distant from the speaker, i.e., casual acquaintances, such as 'the sacristan', or those not personally known, such as 'the American', in (11b), should disfavor the *se*-marked form. Although there is variation between the marked and the unmarked forms in the same contexts, as shown by the pairs of examples in (11), our hypothesis is that, if *salirse* is more subjective than *salir*, the global tendencies in the data will show that *salirse* is more likely to co-occur with close rather than distant third-person subject referents. It turns out that non-specific subjects, such as 'the individual' or 'one' in (12) (coded as distant), are very difficult to find with the *se*-marked form, which supports our hypothesis.

(11) Factor Group: Relationship to speaker

    a. [Speaking of a son] *¡También* se *me* quiso salir *[de la escuela]!*
También tuve que ponerle otro hast'aquí.
'He also wanted to *leave-SE* [school] on me! I had to tell him
again to straighten up.' (Habla popular, 104)
[Speaking of a son] *¿Cuántos años tiene en ese taller? Y de ai no*
ha salido.
'How long has he been in that workshop? And from there he
hasn't *left-0*.' (Habla popular, 406)

    b. <u>*El sacristán*</u> *era un señor ya grande, y fue tal su susto . . . gritaba*
desesperadamente a todos: ". . . ¡Salgan pronto, salgan pronto!"
Y ya él salió *corriendo, y tras él todos nosotros.*
'<u>The sacristan</u> was an older man, and he was so frightened . . .
he yelled desperately to everyone: ". . . Get out fast, get out
fast!" And then he *left-0* running, and behind him all of us.'
(Habla culta, 124)

       Y resulta de que <u>*el americano*</u>, *al bajarnos ahí, en Pino Suárez*
. . . otro señor adentro, en el Metro . . . no lo dejaba bajar. En-
tonces . . . al otro le dio un aventón, y él se salió. *Pero no lo de-*
jaban bajar.
'And it ended up that <u>the American</u>, when we got off there, at
Pino Suárez . . . another man inside, in the metro . . . wasn't let-
ting him get off. Then . . . he gave the other one a push, and he
came out-SE. But they wouldn't let him get off.'
(Habla popular, 123)

(12) *Vamos a suponer que* <u>*el individuo*</u> sale *huyendo; no tiene tiempo . . .*
'Let's suppose that <u>the individual</u> *goes out-0* fleeing; he doesn't have
time . . .' (Habla culta, 399)
Cuando <u>*uno*</u> sale *de la escuela, pues realmente* <u>*uno*</u> sale *a probar –*
¿verdad? –, a probar campos
'When <u>one</u> *leaves-0* school, well, really, <u>one</u> is *going out-0* to try out
[different fields] – right? –, to try out fields' (Habla culta, 33)

Two other factor groups, construction (co-occurrence with preposition *de*
'from') and tense-mood-aspect (Preterit vs. others), attempt to measure
the focus on the change of state and dynamic construal attributed to *se*
in the energetic constructions analysis (Maldonado 1999: 353–373). First,
we hypothesize that *salirse* should be favored in constructions with the
preposition *de* 'from', as *se* is said to profile a particular point in space,
in this case the point of origin as opposed to the entire trajectory of mo-
tion (Maldonado 1999: 367). In the tense-mood-aspect factor group, the
marked form should be favored by perfective aspect coded in the Preterit,

given the association between perfectives and dynamic predicates (e.g., Bybee et al. 1994: 92).

The last factor group considered is polarity. If negation involves ''discourse presupposition'' in which the proposition is familiar or part of shared background (Givón 1984: 328), then the expression of counter-expectation is somewhat incongruous with negation. Earlier example (5c) provides an illustration: When the speaker would 'go out' (*se*-marked) with her friends, she was told she should <u>not</u> 'go out' (0-marked). Thus, we hypothesize that *salirse* rates will be higher in affirmative than in negative polarity contexts.

Multivariate statistical analysis educes regularities and tendencies in the data. It helps to ascertain the effects on speakers' choices of the factors constituting the environment in which each alternate form occurs. The multiple-regression procedure in ''variable-rule (VARBRUL) analysis'', which despite its name does not necessarily involve rules nor impose assumptions about underlying forms, considers simultaneously all factors (contextual elements), both linguistic and social, which are hypothesized to explain the variation (Paolillo 2002; Sankoff 1988b).

Table 1 displays the results for 274 *salir(se)* tokens using the variable-rule analysis application for the Macintosh GOLDVARB 2.1 (Rand and Sankoff 1990). Included were the five linguistic factor groups presented above, each of which corresponds to a hypothesis operationalizing subjectivity or particular claims about the function of *se*-marking, as well as one social factor group, corpus, which tests whether rates of *salirse* are higher in popular (*Habla popular*) than in educated (*Habla culta*) speech. Variable-rule analysis a) identifies the factor groups that contribute a statistically significant effect; b) shows their relative magnitude of effect, by the *Range*; and c) indicates the direction of effect, by assigning *Probability* or *factor weights* to factors between 0 and 1. The closer the values are to 1 the more likely is the *se*-marked form. Conversely, values below .5 favor unmarked *salir* (cf. Poplack and Tagliamonte 2001: 92–94).

The overall likelihood (corrected mean) for the *se*-marked form is .27, corresponding to the overall rate of *salirse* relative to *salir*, at 24% (in the variable context). Shown within brackets are the two factor groups not selected as significant: construction (co-occurring *de*) and corpus (popular vs. educated speech). Given in boldface in the first column are the statistically significant ($p = 0.043$) factor groups: grammatical person, polarity, tense-mood-aspect, and relationship to speaker. These have a similar magnitude of effect, as indicated by Ranges of 25, 24, 23, and 20, respectively. As predicted, the *se*-marked form is favored by first-person singular subjects (with a Probability weight of .68) and third-person referents

Table 1. *Factors contributing to the occurrence of* salir<u>se</u> *as opposed to* salir *in Mexican Spanish (significant factors in boldface; [ ] = not selected as significant); Total N = 274,* p *= 0.043, Corrected mean .27 (24%* salirse*)*

	Probability[a]	% se	Total N	% data
Person				
1st singular	**.68**	32%	76	28%
Other	**.43**	21%	194	71%
Range	25			
Polarity				
Affirmative	**.53**	26%	238	86%
Negative	**.29**	11%	36	13%
Range	24			
Tense-mood-aspect				
Past	**.65**	36%	94	36%
Non-past	**.42**	17%	167	63%
Range	23			
Relationship to speaker				
Close	**.62**	33%	48	40%
Distant/non-specific	**.42**	16%	72	60%
Range	20			
Construction				
+ *de*	[.60]	34%	41	15%
Other	[.48]	22%	227	84%
	n.s.			
Corpus				
Habla popular	[.54]	27%	170	65%
Habla culta	[.43]	17%	88	34%
	n.s.			

[a] Probability (factor) weights above .50 indicate contexts with higher than average *salirse* rates; conversely, values below .50 indicate disfavoring factors.

close to the speaker (.62). Both of these factors provide a measure of speaker involvement.

   The quantitative analyses, however, bring to light some unexpected findings. Recall that perfective aspect was our operationalization of the notion of dynamicity in investigating Maldonado's (1999) energetic constructions analysis. The tense-mood-aspect results show that the relevant distinction is not aspectual, but between past and non-past temporal reference. Rather than singling out perfective aspect as hypothesized, both Preterit (past perfective) and Imperfect (past imperfective) combined favor *salirse* (Probability .65). Table 2 shows virtually identical *salirse* rates in the Preterit, at 36% (26/60), and the Imperfect, at 35% (12/34). In contrast, Present forms show a lower than average *salirse* rate, at 21% (21/98), while infinitive forms nearly categorically eschew the *se*-marked form, at 2% (1/38).

Table 2.   Salir*se*/salir *tense distribution*

	Salirse	*Salir*	Total N
Preterit	36%	64%	60
Imperfect	35%	65%	34
Present	21%	79%	98
Infinitive	2%	98%	38

Why is the infinitive so unfavorable to *salirse*? We propose it is because of the "irrealis" character of the infinitive (cf. Haspelmath 1989: 288). Even when exiting is against an obstacle, as in (13a) (*el aguacero* 'the downpour'), or for pleasure, as in (13b) (*me gusta mucho* 'I like [it] a lot'), it is not marked by *se* if it is not (yet) realized.

(13)   a.   *Y el aguacero en toda su fuerza, que era imposible poder* salir *de aquel carro. Por fin, era tal nuestra angustia, que con todo y el aguacero nos decidimos a* salir *del carro.*
   'And the downpour at full force, it was impossible to *get out-0* of that car. Finally, we were so distressed, that with the downpour and everything we decided to *leave-0* the car.'

(México, Habla culta, 128)

   b.   *Y a mí si me gusta mucho* salir. *Y ora le digo: "Yo pensé que no m'iban a dejar ir a l'escursión".*
   'And I really like to *go out-0*. And so now I tell him: "I thought you weren't going to let me go on the trip".'

(México, Habla popular, 245)

Another unexpected result is the lack of significance for co-occurrence with preposition *de* 'from', which tested *se*'s focusing function as profiling the origin of motion (Maldonado 1999: 367). Though indicating a direction of effect in the predicted direction (at 34%), it does not achieve statistical significance when considered simultaneously with the other factor groups. This result, together with the absence of a perfectivity effect, fails to provide support for a characterization of *salirse* in terms of a dynamic abrupt or rapid event as proposed in the energetic constructions analysis (Maldonado 1999: 369). Rather, an account in terms of realis-irrealis (realized vs. non-realized situation) may be more applicable in characterizing actual usage, at least in the present data.

A characterization of *salirse-salir* in terms of realis-irrealis is also consonant with the polarity factor group results, where negative polarity shows a disfavoring effect (Probability .29). Thus, going out for pleasure is not *se*-marked if not realized (14). (Note that these are

aggregate tendencies for a variable phenomenon: Though it is not favored, *salirse* does sometimes appear under negation, as in [15].) More importantly, the strong polarity effect supplies nice empirical confirmation for a counter-expectation meaning since negation involves presupposition and hence (at least some degree of) anticipation.

(14) *Que ya <u>no</u> salgo a alguna fiesta; ya no voy*
      'I <u>don't</u> go out-0 to parties anymore; I don't go anymore'
                                                    (Habla popular, 80)

(15) *Dice: "Pues se me sale inmediatamente." Le digo: "No. <u>No</u> me puedo salir, porque yo soy vocal, y es muy importante – le digo . . ."*
      'He says: "Well get out of here immediately". I tell him: "No. I ca<u>n't</u> leave-SE, because I'm a member, and it's very important", I tell him . . .'
                                                    (Habla popular, 109)

In summary, we have shown how accounting for variable marking of a grammatical morpheme reveals patterns of subjectivity. For middle marker *se* on intransitive motion verb *salir*, subjectivity is manifested structurally in the tendency to occur with a first-person singular subject (the speaker) or a third-person subject close to the speaker, and in the past tense (when the event is already realized), and conversely, in its tendency to shun negative polarity (presupposition) contexts. Thus, the probability weights yielded by the analysis of variable use of a grammatical morpheme enable us to identify measures of its subjectification.

5. Dialectal and diachronic evidence for subjectification

Recall that *salirse* rates are not significantly higher in construction with the preposition *de* (Table 1). This unpredicted finding points us toward what turns out to be an important locus of change in the subjectification of *salirse*. We will present data that suggest that *salirse* has had a counter-expectation meaning since its earliest uses and that semantic change lies in increasingly speaker-based construal of expectation. This change occurs as contextual meaning in the previously highly frequent *de* construction is absorbed into the *salirse* form itself.

Dialect differences may reflect different stages of evolution (Silva-Corvalán 2001: 16). The comparison of *salirse* use in different varieties of Spanish, shown in Table 3, turns out to be quite revealing. In the Mexican corpora, token frequency normalized per 100,000 words, at 19, and frequency relative to *salir*, at 12% (65/565), are four to five times greater than in the Peninsular (= Spain) data, where *salirse* is quite scarce, with a

Table 3.   *Comparison of* salir<u>se</u> *frequencies in Mexican, Peninsular, and Old Spanish*

Variety (word count; corpus)	Normalized token frequency (per 100,000 words)	Relative frequency[a] (*salirse / salir*)
Mexican Spanish (~340,000; Habla culta, popular)	19	12% (65/565)
Peninsular Spanish (~380,000; Habla culta, COREC)	3	3% (13/486)
Old Spanish (12th to 15th c.) (~1,110,000; see Appendix)	4	4% (43/1163)

[a] Differences Mexican–Peninsular, Mexican–Old Spanish significant; difference Peninsular–Old Spanish n.s. (Chi-square tests).

normalized frequency of 3 and a relative frequency of 3% (13/486)—only 13 tokens in close to 400,000 words of text. In Old Spanish (12th to 15th century) texts, the *se*-marked form has a normalized frequency of 4 and a relative frequency of 4% (43/1163). That is, frequencies of *salirse* in Old Spanish are three to four times lower than in the Mexican corpora, but comparable to present-day Peninsular frequencies.

Higher *salirse* frequencies in the Mexican data may be taken as an indicator of change (cf. Croft 2000: 57). What kind of semantic change is involved? To answer this question, let us look at some of the rare Peninsular tokens. One *salirse* context is that of escaping, for example, the rabbits that ate up the sack in (16). Another use is to physically get out of line. For inanimate subjects, this may mean 'protrude, stick out, spill over', as with the blanket in (17), while for humans it applies to driving off the road, as with the racecar accident in (18). *Salirse* is used for ailing or injured body parts, for example, *el codo se me ha salido* 'my elbow popped out' (COREC, CCON004D), parallel to Mexican examples with datives (see [10]). There was one case of moving out of home, shown in (19), but none of *salirse* from an organization or job. Note that the Peninsular *salirse* occurrences are mostly in "force dynamics" (Talmy 1985) situations of a physical nature: exiting from a physically delimited space or against a physical obstacle (the road, the sack). We did not find examples of obstacles or opposition of a more abstract character, such as socio-cultural norms, which abound in the Mexican data.

(16)   Peninsular *salirse* 'to escape'
[Talking about hunting rabbits] *¿Te acuerdas que <u>nos comieron el saco</u> y* se *nos* salieron *unos pocos?*
'Do you remember <u>they ate the sack on us</u> and a few of them *got away-SE* from us?'   (COREC, CCON019A)

(17) Peninsular *sal<u>ir</u>se* 'get out of line, protrude, spill over'
[Talking about blankets]
no sé yo si no arrastrará mucho, yendo de noventa para la cama de ochenta . . .
No, sólo que se remete más.
Está bien porque queda más agregada, ¿no?
Claro. No te, no se te, no se *te* sale.
'I don't know if it won't trail a lot, going from [size] ninety for the [size] eighty bed . . .
No, you just have to tuck it in more.
It's okay because it stays tighter, right?
Of course. It doesn't, it doesn't, it doesn't *get out of place-SE* on you.'   (COREC, CCON013C)

(18) Peninsular *sal<u>ir</u>se* 'drive off road'
Mira, hablando de sidecars, ayer hubo el único accidente grave que se ha registrado en los entrenamientos . . . un piloto . . . – pues, bueno, al final de la recta se salió –
'Look, speaking of sidecars, yesterday was the only serious accident that has occurred in the practices . . . a driver . . . – so, well, at the end of the racecourse he *went off-SE* [the course] –'
(COREC, ECON006B)

(19) *Alegando . . . "la casa es mía, preferiría que os marcharais vosotros".*
Bien.
Y ¡qué le vamos a hacer!
Pues . . . ya que casi lleváis la razón, nos salimos – *je, je!*
Nos salimos.
'Claiming . . . "the house is mine, I would prefer it if you were the ones to leave". Fine.
And there's nothing to be done about it!
Well . . . since you're almost right, we'll *leave-SE*, ha ha!
We'll *leave-SE*.'   (Madrid, 333)

Now, if Peninsular Spanish represents a less advanced stage in the semantic evolution of *salirse*, these examples suggest that, in an earlier stage, the marked form used to appear more in contexts of physically getting out of line or deviating from a delimited space. This would be as expected since the directionality of meaning change is commonly from concrete to abstract (cf. Hopper and Traugott 1993: 77 and references therein). Diachronic data provide support for such an earlier stage. The following examples show that *salirse* uses in Old Spanish texts are recognizable as ones that persist today. Common is 'escape', usually against a physical

obstacle, as with the lion escaping from the net or the ''prisoner'' slipping through the speaker's hands (20), though we also find cases of abstract obstacles, when exiting against someone else's wishes or in an emotionally difficult situation, for example, leaving a room/building in refusal to comply with a request or having failed in a mission (21).

(20)   Old Spanish *salirse* 'to escape (physical)'
   a.   Salios <u>de la Red</u> & desatos el Leon.
        'The lion *got out-SE* <u>from the net</u> and untied himself.'
        (12th c., Cid, verse 2282)
   b.   *mio preso es e yo lo deuo soltar quando me yo quesiere; e non querria que* se *me* saliese <u>de manos</u> *por alguna maestria*
        'he is mine and I should let him go when I would want to; and I wouldn't want him to *slip-SE* <u>from my hands</u> by some trick'
        (14th c., Zifar, 66)

(21)   Old Spanish *salirse* 'to ''escape'' (abstract)'
   a.   *nj<n> quiso fazer . . . lo q<ue> ella demandaua. &' come<n>c'os a* salir *dela camara por yr se. % Esto<n>ces ella q<u>a'ndo esto uio. echol mano enel ma<n>to*
        'he didn't want to do . . . what she was demanding and he started to *go out-SE* of the room to leave. Then when she saw this, she grabbed his mantle'   (13th c., GE1)
   b.   *despedime de ella con más lágrimas que palabras, y después de besarle las manos* salime *de palacio con un nudo en la garganta*
        'I bade her farewell with more tears than words, and after kissing her hands I *left-SE* the palace with a lump in my throat'
        (15th c., Cárcel, el autor, parte 5)

Another set of Old Spanish occurrences involves driving off the road as in *salirse de la carrera, del camino* 'run off the road' (22). The final group of Old Spanish examples is of standing apart, separating from, permanently leaving a group or locale (23). (Naturally, there is variation between the marked and unmarked forms, as illustrated in (24): he left-*se*, but he ordered that the men leave-0.)

(22)   Old Spanish *salir*se 'get/drive off road'
   a.   *&' co<n> miedo q<ue> ouo ell asna del* salios <u>d<e>la carrera.</u> *& come<n>c'os a yr por defuera por un campo*
        'and with the fear the donkey felt toward him, it *went off-SE* <u>(from) the road</u>. And it started to go on the outside through a field'   (13th c., GE1)

b. *Mientras el escudero ... cabalgaba adormecido, el palafrén,* sa-
liéndose *de camino, se metió en medio del bosque, divagando a
su antojo*
'While the squire ... rode along asleep, the horse, *going off-SE
(from) the road*, went into the middle of the forest, meandering
wherever it would'   (13th c., Caballería, prólogo)

(23) Old Spanish *salir*se 'leave/abandon a group', 'leave permanently'
a. *destruye el caballero en sí mismo la caballería cuando desama el
oficio de caballero, o* se sale *del orden de caballería.*
'the knight destroys chivalry when he disrespects the occupa-
tion of knight, or when he *leaves-SE* (from) the knighthood'
(13th c., Caballería, parte segunda)
b. *pharaon ma<n>do a abraha<m> q<ue>* se saliesse *daq<ue>lla
tierra*
'the Pharaoh ordered Abraham to *leave-SE* (from) that land'
(13th c., GE1)

(24) *E por q<ue> se no<n> fiziesse<n> y mas Nin<n>os ... partio los
uarones de las mug<ie>r<e>s. E tomo el su conpan<n>a &* salio se
*fuera dela villa co<n> ella. & mando a todos los' uarones dela cibdat
& dela tierra q<ue>* saliesse<n> *& fuessen co<n> el luego alli. de
morada*
'And so that they wouldn't make any more children ... he sepa-
rated the men from the women, and he took his company and he
left-SE the village with it, and he ordered all of the men of the city
and of the land to *leave-0* and to go with him then to live there'
(13th c., GE1)

Why might 'escaping' and 'driving off the road' be (relatively) common
Old Spanish *salirse* uses? It seems that the notion shared by these contexts
of use is that of 'spilling', which is recurrent in early dictionary entries for
salirse. For example, the 1739 *Autoridades* dictionary definition is "caerse
lo que está contenido en otra cosa por alguna rotúra" [to fall out that
which is contained in something else through some rupture] (the second
part of the definition is "apartarse, ò echarse fuera de lo contratado, ò
pactado" [separate from, or go back on something agreed upon], as in
[23]/[24]) (RAE 1739: 24). This testimony, together with early examples
such as *salirseme el alma* 'my soul leaving me' (15th c., *Celestina*, 178),
suggests that *salirse* originates with a 'spill over' or 'burst through' mean-
ing which is manifested for animate subjects in escaping and driving off
the road. In summary, the Old Spanish examples confirm the suggestion
from present-day Peninsular Spanish that *salirse* originates in contexts of

physically deviating from a delimited space or getting out of line. We did find more abstract Old Spanish cases of exiting in the face of another's will or leaving the site of unpleasantness (21), but no cases of going out against personally construed or social norms, or to have a good time, as in the Mexican data.

We can now address the nature of the subjectification process of *salirse* as a counter-expectation marker. Drawing on the notion of "force dynamics" (Talmy 1985), or the relationship of barriers, we have distinguished three kinds of 'obstacles' to exiting expressed by *salirse*:

(i)   physical obstacles, such as containers, nets, roads, fences [cf. (16), (18), (20), (22)];

(ii) more abstract but individual obstacles, such as specific rules and another's will [cf. (2a), (2c), (21a)];

(iii) general social norms, such as the idea that there are appropriate occasions and manners of exiting or going out [cf. (1a), (5)].

In his energetic constructions analysis, Maldonado (1999: 375) has pointed to "an abstract confrontation of force dynamics ... [N]atural expectations regarding different world events constitute the initial force that a particular event confronts" [our translation]. In our view, however, speaker expectations come into play with all three kinds of obstacles, from the earliest examples of *salirse*. (Even with a lion or rabbits physically breaking out of a net or sack [20a], [16], the speaker has an expectation/desire about that event.) Thus, we would like to take Maldonado's appeal to abstract force dynamics one step further. What changes is that the obstacles to *salirse* become less objectively apparent and more speaker-construed, in other words, the speaker's perspective takes on an increased weight. Thus, increasingly speaker-based meaning (Traugott and Dasher 2002: 94–96) in the subjectification of *salirse* lies in the increasingly speaker-based construal of (counter-)expectation, as the situations come to include abstract force dynamics ENCOMPASSING SOCIAL NORMS. Thus, the locus of the subjectification process is speaker construal of expectation.

Empirical support for the diachronic subjectification process we propose appears in a comparison of the distribution of the *salirse* form in Old and Mexican Spanish, shown in Table 4. Co-occurring datives show no change, thus a higher level of affectedness (Maldonado 1999: 394) may have been a component of *salirse* meaning since the beginning. The proportion of past tense (Preterit and Imperfect) occurrences also remains unchanged, though there may have been a decrease in the Preterit (in the present data, from 51% to 34%), which could indicate weakening over time of an earlier aspectual constraint. That is, dynamicity as understood

Table 4.   *Changes from Old Spanish to present-day Mexican Spanish* salirse

	Old Spanish (N = 45)	Mexican Spanish (N = 66)
Co-occurring datives	18%	15%
Past tense	53%	52%
Preterit	51%	34%
Inanimate subject[a]	**29%**	**14%**
First-person singular[a]	**11%**	**38%**
Co-occurring ***de***	**58%**	**21%**

[a] In Peninsular data, inanimate subjects at 38% (5/13), first-person singular at 15% (2/13). Difference between combined Old Spanish/Peninsular and Mexican proportion of inanimate subjects is significant (Chi-square 5.236667086, $p = 0.0221$).

in an ''energetic constructions'' analysis (Maldonado 1999) may indeed have been a significant factor in *salir(se)* variation at an earlier stage of development.

Where there is evident change is in *salirse* subjects. Inanimate subjects decline from 29% in Old Spanish to 14% in Mexican, while first-person singular subjects increase more than threefold, from 11% to 38%. These differences cannot be attributed to genre differences alone. The (admittedly meager) Peninsular data also show a higher proportion of inanimates, at 38% (5/13), and a lower proportion of first-person singular, at 15%, even though, like the Mexican and unlike the Old Spanish, they are conversational oral data. Their higher proportion in earlier stages supports our conjecture of an important role for inanimate subjects in the origins of *salirse*.

The most striking difference between the Old Spanish and Mexican data shows up in co-occurring *de*, already illustrated in several examples, cf. (20)–(23). The proportion of co-occurring *de* drops from 58%, more than half, to 21%, about one-fifth (Table 4). This change provides empirical support for Maldonado's (1999: 390) proposal for an earlier focusing function (focus on the change of state at the origin of motion), which we take to be indicated by high rates of co-occurring *de* 'from' in the Old Spanish data. More importantly for our purposes, the contrasting decreased rate of construction with *de* in the present-day Mexican data serves as an indication of increasingly non-physical, abstract situations. That is, the shift from the earlier preponderance of co-occurring *de* indicates a shift from concrete, exterior spatial limits or obstacles to abstract ones dependant upon the speaker's evaluation, as '*salirse*' from nets and sacks extends to '*salirse*' beyond personally construed or social bounds.

The broader theoretical import of the shift in *de* co-occurrence is the evidence provided for absorption of contextual meaning as a mechanism

of semantic change. Contextual absorption differs somewhat from pragmatic inference (Hopper and Traugott 1993: 63–93) in that it is the linguistic rather than the pragmatic context that is conventionalized (Bybee et al. 1994: 296). For example, in the grammaticization of the Spanish Present Perfect as a hodiernal perfective (past situations occurring on the same day), the form "absorbed the contextual meaning" accompanying frequently co-occurring adverbial modification: "the frequent use of the P[resent] P[erfect] with 'today' temporal adverbs led to the inference that 'today' was actually part of the meaning of the PP gram" (Schwenter 1994: 102). In our case, early frequent construction with *de* and its subsequent drop between Old and Mexican Spanish suggests that the focusing function of the co-occurring *de* construction is similarly absorbed into the meaning of *salirse*. The semantic change from exiting against PHYSICAL OBSTACLES to exiting against SOCIAL NORMS thus appears to proceed via absorption of contextual meaning. The 'outside-the-norm' meaning has been conventionalized in the familiar expression in colloquial Mexican Spanish for indicating exaggerated or out of line behavior, *te sales*, which is fixed in its occurrence without an accompanying *de* phrase (25).

(25)   *te   sales*
       2SG-REFL   go.out-PRES-2S
       'you're out of line'

6.   Conclusion

This study has employed quantitative methods to bring forward a set of findings on the subjectivity of middle *se*-marked *salirse* synchronically and on the subjectification of this form diachronically. Analysis of synchronic variation between marked and unmarked *salir(se)* in present-day Mexican Spanish shows that the probability of occurrence of the *se*-marked form is greatly reduced in irrealis (infinitives) and negative polarity contexts, but favored by the past tense as well as by first-person singular and third-person subjects referring to persons close to the speaker. Dialectal and diachronic evidence for subjectification emerges from two changes. The first is the decreased proportion of inanimate subjects and correspondingly increased association with human subjects, in particular the speaker. This distribution change points to origins with a 'spill over' or 'burst through' meaning and extension to human referents going against a natural order of events. The second change is the drop in co-occurring *de* 'from', which points to the absorption of the 'exit from' meaning of the frequent co-occurring *de* construction into the *salirse* form, in tandem with out-of-line or beyond-the-bounds uses. A broader implication for

our understanding of the subjectification of counter-expectation markers is that increased "expressiveness" (Traugott and Dasher 2002: 94–96) in the case of *salirse* lies in the increasingly speaker-based construal of (counter-)expectation, as the use of the form extends to situations involving non-physical abstract, even social, force dynamics.

More generally, we have shown how an account of variation in the use of a grammatical morpheme can help identify the contextual factors that constitute operational measures of subjectification. While some factors, such as first-person singular, have been foreseen in previous work on subjectification, others, namely referent relationship to speaker and polarity, emerge as structural correlates of subjectification from the quantitative analyses here for the first time. It is possible that cross-linguistic regularities in the distribution patterns evincing subjectification will be particular to functional domains, such as the middle voice, (future) tense, or clause connectives. How general is the applicability of the subjectification measures found in this study? More empirical case studies measuring subjectification will tell.

Appendix

Corpora are listed in chronological order; word count and *salirse/salir* frequency appear within [ ].

Cid (ca. 1140) = Menéndez Pidal, Ramón, ed. 1944. *Cantar de mio Cid. Texto, gramática y vocabulario*, vol. 3: *Texto*. Madrid: Espasa Calpe. [30,000, 3/65]

Fazienda (12th c.) = Almerich, Arcediano de Antiochia. 1965. *La Fazienda de Ultra Mar. Biblia Romanceada et Itinéraire Biblique en prose castillane du XIIe siécle*, ed. Moshé Lazar. In *Acta Salmanticensia, Filosofía y letras*, vol. 18, no. 2. Salamanca: Universidad de Salamanca. [85,500, 3/73]

Calila (1250) = Anónimo. 1987. *Calila e Dimna*, ed. J. M. Cacho Blecua and M. J. Lacarra. Madrid: Castalia. [86,000, 4/75]

GE1 (1272–1275) = Alfonso X. 1997. *General estoria I*: Madrid: Nacional 816. On CD-Rom: *The Electronic Texts and Concordances of the Prose Works of Alfonso X, el sabio*, prepared by Lloyd Kasten, John Nitti, and Wilhelmina Jonxis-Henkemans. Madison: Hispanic Seminary of Medieval Studies. [577,000, 16/613]

Apolonio (13th c.) = Anónimo. *Libro de Apolonio*, códice III–K–4 de la Biblioteca del Escorial, based on the edition by Dolores Corbella (Libro de Apolonio, Madrid, Cátedra, 1999). Available at: Biblioteca Virtual Miguel de Cervantes, http://www.cervantesvirtual.com/. [20,000, 1/45]

Caballería (13th c.) = Lulio, Raimundo. *Libro del orden de caballería; Príncipes y juglares*. Available at: Biblioteca Virtual Miguel de Cervantes, http://www.cervantesvirtual.com/. [25,000, 4/7]

Engaños (13th c.) = Anónimo. 1946. *Libro de los Engaños*, edición de Ángel González Palencia. Madrid: Versiones castellanas del Sendebar, C. S. I. C. Available at: Biblioteca Virtual Miguel de Cervantes, http://www.cervantesvirtual.com/. [15,000, 2/21]

Zifar (1320) = Anónimo. 1929. *El libro del Cavallero Zifar*, edición de Ch. Ph. Wagner, Ann Arbor: University of Michigan Press. [33,500, 2/56]

Lucanor (1350) = Don Juan Manuel. 1971. *El conde Lucanor o Libro de los enxiemplos del conde Lucanor et de Petronio*, ed. J. M. Blecua. Madrid: Castalia. [74,000, 0/51]

Corbacho (1438) = Martínez de Toledo, Alfonso. 1978. *Corbacho*, ed. M. Gerli. Madrid: Cátedra. [75,000, 0/69]

Cárcel (1492) = de San Pedro, Diego. 1972. *Cárcel de Amor*, ed. Keith Whinnom, Madrid: Castalia. [25,500, 2/24]

Celestina (1499) = de Rojas, Fernando. 1987. *La Celestina*, ed. D. S. Severin. Madrid: Cátedra. [67,000, 6/64]

Habla Culta = Lope Blanch, Juan M., ed. 1971. *El habla de la ciudad de México: materiales para su estudio*. México: UNAM. [167,000, 15/237]

Habla Popular = Lope Blanch, Juan M., ed. 1976. *El habla popular de México: materiales para su estudio*. México: UNAM. [172,000, 47/328]

Madrid = Esgueva, E. and M. Cantarero, eds. 1981. *El habla de la ciudad de Madrid: materiales para su estudio*. Madrid: CSIC. [140,000, 6/216]

COREC = Corpus de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular, director Francisco Marcos Marín. Available at: www.lllf.uam.es/~fmarcos/informes/corpus/corpusix.html (género conversacional). [241,000, 7/270]

Notes

1.   Examples are reproduced exactly from the corpus cited, except for omitted material indicated by … and context summarized within [ ].

2. In these data, in 'permanently leaving a job' contexts, *salirse* occurs when quitting, *salir* when retiring or with negative polarity, for example, *¿Cuántos años tiene en ese taller? Y de ai no ha* salido 'How many years have you been in that workshop? And you haven't *left-0* there' (Habla popular, 406).

References

Aaron, Jessi Elana
  2004   The gendered use of *salirse* in Mexican Spanish: *Si me salía yo con las amigas, se enojaba. Language in Society* 33, 585–607.
Barlow, Michael and Suzanne Kemmer (eds.)
  2000   *Usage-based Models of Language.* Stanford, CA: CSLI Publications.
Benveniste, Émile
  1966   De la subjectivité dans le langage. In Benveniste, Émile, *Problèmes de linguistique générale.* Paris: Gallimard, 258–266.
Butt, John and Carmen Benjamin
  2000   *A New Reference Grammar of Modern Spanish*, 3rd ed. Lincolnwood, IL: McGraw Hill.
Bybee, Joan
  2005   The impact of usage on representation: Grammar is usage and usage is grammar. Presidential address, annual meeting of the Linguistic Society of America, Oakland, CA.
Bybee, Joan and Paul Hopper (eds.)
  2001   *Frequency and the Emergence of Linguistic Structure.* Amsterdam/ Philadelphia: Benjamins.
Bybee, Joan, Revere Perkins, and William Pagliuca
  1994   *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World.* Chicago, IL: The University of Chicago Press.
Company Company, Concepción
  to appear   Zero in syntax, ten in pragmatics, or subjectification as syntactic cancellation. In Canakis, Costas, Sophia Athanasidou, and Bert Cornillie (eds.), *New Pathways of Subjectivity.* Berlin/New York: Mouton de Gruyter.
Croft, William
  2000   *Explaining Language Change. An Evolutionary Approach.* London: Longman.
Diver, William
  1985   The history of linguistics in the west: How the study of language went wrong in the western tradition. Transcribed by Takashi Aoyama. Tezukayama College, Japan: *Special English Linguistics Series*, 43–68.
Dubois, John W.
  1987   The discourse basis of ergativity. *Language* 63, 805–855.
Finegan, Edward
  1995   Subjectivity and subjectivisation: An introduction. In Stein, Dieter and Susan Wright (eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives.* Cambridge: Cambridge University Press, 1–15.
Givón, Talmy
  1984   *Syntax. A Functional-Typological Introduction*, vol. 1. Amsterdam/ Philadelphia: Benjamins.
Haspelmath, Martin
  1989   From purposive to infinitive—A universal path of grammaticization. *Folia Linguistica Historica* 10 (1/2), 287–310.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer
  1991   *Grammaticalization: A Conceptual Framework*. Chicago, IL: University of Chicago Press.
Hopper, Paul J. and Elizabeth Closs Traugott
  1993   *Grammaticalization*. Cambridge: Cambridge University Press.
Kemmer, Suzanne
  1993   *The Middle Voice*. Amsterdam/Philadelphia: Benjamins.
Jakobson, Roman
  1960   Concluding statement: Linguistics and poetics. In Sebeok, T. A. (ed.), *Style in Language*. New York: John Wiley and Sons, 350–377.
Labov, William
  1969   Contraction, deletion, and inherent variability of the English copula. *Language* 45 (4), 715–762.
  1972   *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.
Langacker, Ronald W.
  1990   Subjectification. *Cognitive Linguistics* 1, 5–38.
  1991   *Foundations of Cognitive Grammar. Vol. 2: Descriptive Applications*. Stanford, CA: Stanford University Press.
Maldonado, Ricardo
  1999   *A media voz: Problemas conceptuales del clítico* se. México: UNAM.
Milroy, Leslie and Matthew Gordon
  2003   *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
Paolillo, John
  2002   *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford, CA: CSLI Publications.
Poplack, Shana and Sali Tagliamonte
  1999   The grammaticization of *going to* in (African American) English. *Language Variation and Change* 11, 315–342.
  2001   *African American English in the Diaspora*. Oxford: Blackwell.
Rand, David and David Sankoff
  1990   GoldVarb 2.1: A Variable Rule Application for the Macintosh. Montréal: Centre de recherches mathématiques, Université de Montréal. Available at: http://www.crm.umontreal.ca/~rand/GoldVarb_Eng.html.
Real Academia Española
  1739   *Diccionario de la lengua castellana, en que se explica el verdadero sentido de las voces, su naturaleza y calidad, con las phrases o modos de hablar, los proverbios o refranes, y otras cosas convenientes al uso de la lengua [...]. Compuesto por la Real Academia Española. Tomo sexto. Que contiene las letras S.T.V.X.Y.Z.* Madrid: Imprenta de la Real Academia Española, por los herederos de Francisco del Hierro. Reproducido a partir del ejemplar de la Biblioteca de la Real Academia Española. Available at: http://buscon.rae.es/ntlle/SrvltGUILoginNtlle.
Sankoff, David
  1988a   Sociolinguistics and syntactic variation. In Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge Survey, Vol. IV*. Cambridge: Cambridge University Press, 140–161.
  1988b   Variable rules. In Ammon, Ulrich, Norbert Dittmar, and Klaus J. Mattheier (eds.), *Sociolinguistics: An International Handbook of the Science of Language and Society*. Berlin/New York: Walter de Gruyter, 984–997.

Scheibman, Joanne
  2002   *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation.* Amsterdam/Philadelphia: Benjamins.

Schwenter, Scott A.
  1994   The grammaticalization of an anterior in progress: Evidence from a peninsular Spanish dialect. *Studies in Language* 18, 71–111.

Schwenter, Scott A. and Elizabeth Closs Traugott
  1995   The semantic and pragmatic development of substitutive complex prepositions in English. In Jucker, Andreas H. (ed.), *Historical Pragmatics: Pragmatic Developments in the History of English.* Amsterdam/Philadelphia: Benjamins, 244–273.

Silva-Corvalán, Carmen
  1994   *Language Contact and Change: Spanish in Los Angeles.* Oxford: Clarendon Press.
  2001   *Sociolingüística y pragmática del español.* Washington, D.C.: Georgetown University Press.

Stein, Dieter and Susan Wright (eds.)
  1995   *Subjectivity and Subjectivisation: Linguistic Perspectives.* Cambridge: Cambridge University Press.

Talmy, Leonard
  1985   Force dynamics in language and cognition. *Cognitive Science* 2, 49–100.

Torres Cacoullos, Rena and Scott A. Schwenter
  2005   Towards an operational notion of subjectification. *Proceedings of the Thirty-First Annual Meeting of the Berkeley Linguistics Society.*

Traugott, Elizabeth Closs
  1989   On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65, 31–55.
  1995   Subjectification in grammaticalization. In Stein, Dieter and Susan Wright (eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives.* Cambridge: Cambridge University Press, 31–54.
  1999   The rhetoric of counter-expectation in semantic change: A study in subjectification. In Blank, Andreas and Peter Koch (eds.), *Historical Semantics and Cognition.* Berlin/New York: Mouton de Gruyter, 177–196.
  2003   Constructions in grammaticalization. In Joseph, Brian D. and Richard D. Janda (eds.), *The Handbook of Historical Linguistics.* Oxford: Blackwell, 624–647.

Traugott, Elizabeth Closs and Richard B. Dasher
  2002   *Regularity in Semantic Change.* Cambridge: Cambridge University Press.

Verhagen, Arie
  1995   Subjectification, syntax, and communication. In Stein, Dieter and Susan Wright (eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives.* Cambridge: Cambridge University Press, 103–128.